

# 24th Annual National Conference on Managing Environmental Quality Systems

**8:30 – 12:00 TUESDAY, APRIL 12<sup>TH</sup> - A.M. Stockholder Meetings**

**12:00 – 4:30 TUESDAY, APRIL 12<sup>TH</sup>**

**Opening Plenary** (Salons A-H)

- Opening Address
  - Reggie Cheatham, Director, OEI Quality Staff, EPA
  - Linda Travers, Principal Deputy Assistant Administrator, OEI, EPA
- Invited Speakers
  - Tom Huetteman, Deputy Assistant Regional Administrator, EPA Region 9
  - John Robertus, Executive Officer of San Diego Regional Water Quality Control Board, Region 9
- Keynote Address
  - Thomas Redman, President, Navesink Consulting Group
- Panel Sessions
- **Value of the Data Quality Act—Perspectives from OMB, Industry, and EPA (VDQA)**
  - Nancy Beck, OMB
  - Jamie Conrad, American Chemistry Council
  - Reggie Cheatham, Director, OEI Quality Staff, EPA
- **Wadeable Streams: Assessing the Quality of the Nation's Streams (WS)**
  - Margo Hunt, Panel Moderator
  - Mike Shapiro, Deputy Assistant Administrator, Office of Water
  - Steve Paulsen, Research Biologist, ORD

**8:30 – 10:00 WEDNESDAY, APRIL 13<sup>TH</sup>**

**Environmental Measures (EM)** (Salons A-C) *Chair: L. Bradley, EPA*

- Data Error Reduction by Automation throughout the Data Workflow Process (A. Gray, EarthSoft, Inc.)
- Analytical Approaches to Meeting New Notification Levels for Organic Contaminants in Calif. (D. Wijekoon, Calif. DHS)
- Streamlining Data Management and Communications for the Former Walker AFB Project (R. Amano, Lab Data Consultants, Inc.)

**Quality System Implementation in the Great Lakes Program (QSI-GLP)** (Salon D) *Chair: M. Cusanelli, EPA*

- GLNPO's Quality System Implementation for the New "Great Lakes Legacy Act for Sediment Remediation" (L. Blume, EPA)
- Black Lagoon Quality Plan Approval by GLNPO, MDEQ, ERRS, and USACE (J. Doan, Environmental Quality Management, Inc.)
- Remediation of the Black Lagoon Trenton Channel . . . Postdredging Sampling & Residuals Analysis (J. Schofield, CSC)

**Quality Systems Models (QSM)** (Salons F-H) *Chair: G. Johnson, EPA*

- Improving E4 Quality System Effectiveness by Using ISO 9001: 2000 Process Controls (C. Hedin, Shaw Environmental)

**Applications of Novel Techniques to Environmental Problems (ANTEP)** (Salon E) *Chair: B. Nussbaum, EPA*

- On Some Applications of Ranked Set Sampling (B. Sinha, University of Maryland)
- Combining Data from Many Sources to Establish Chromium Emission Standards (N. Neerchal, University of Maryland)
- Estimating Error Rates in EPA Databases for Auditing Purposes (H. Lacayo, Jr., EPA)
- Spatial Population Partitioning Using Voronoi Diagrams For Environmental Data Analysis (A. Singh, UNLV)

**Ambient Air Session I (Sierra 5&6) Chair: M.Papp, EPA**

- Changes and Improvements in the Ambient Air Quality Monitoring Program Quality System (M. Papp, EPA)
- Guidance for a New Era of Ambient Air Monitoring (A. Kelley, Hamilton County DES)
- Environmental Monitoring QA in Indian Country (M. Ronca-Battista, Northern Arizona University)
- Scalable QAPP IT Solution for Air Monitoring Programs (C. Drouin, Lake Environmental Software)

**10:30 – 12:00 WEDNESDAY, APRIL 13<sup>TH</sup>**

**Environmental Laboratory Quality Systems (ELQS) (Salons A-C) Chair: L. Bradley, EPA**

- A Harmonized National Accreditation Standard: The Next Step for INELA Field Activities (D. Thomas, Professional Service Industries, Inc.)
- Development of a Comprehensive Quality Standard for Environmental Laboratory Accreditation (J. Parr, INELA)
- Advanced Tracking of Laboratory PT Performance and Certification Status with Integrated Electronic NELAC-Style Auditing Software (T. Fitzpatrick, Lab Data Consultants, Inc.)

**Performance Metrics (PM) (Salon D) Chair: L. Doucet, EPA**

- Formulating Quality Management Metrics for a State Program in an Environmental Performance Partnership Agreement (P. Mundy, EPA)
- How Good Is “How Good Is?” (Measuring QA) (M. Kantz, EPA)
- Performance-Based Management (J. Santillan, US Air Force)

**Quality Assurance Plan Guidance Initiatives (QAPGI) (Salons F-H) Chair: A. Batterman, EPA**

- A CD-ROM Based QAPP Preparation Tool for Tribes (D. Taylor, EPA)
- Military Munitions Response Program Quality Plans (J. Sikes, U.S. Army)

**Ask a Statistician: Panel Discussion (Salon E) Moderator: B. Nussbaum, EPA Panelists:**

- Mike Flynn, Director, Office of Information Analysis and Access, OEI, EPA
- Reggie Cheatham, Director, Quality Staff, OEI, EPA
- Tom Curran, Chief Information Officer, OAQPS, EPA
- Diane Harris, Quality Office, Region 7, EPA
- Bill Hunt, Visiting Senior Scientist, North Carolina State University (NCSU)
- Rick Linthurst, OIG, EPA

**Ambient Air Session II (Sierra 5&6) Chair: M. Papp, EPA**

- National Air Toxics QA System and Results of the QA Assessment (D. Mikel, EPA)
- Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) of the National Air Toxics Trends Stations (NATTS) and Supporting Laboratories (S. Stetzer Biddle, Battelle)
- Interlaboratory Comparison of Ambient Air Samples (C. Pearson, CARB)
- Developing Criteria for Equivalency Status for Continuous PM<sub>2.5</sub> Samplers (B. Coutant, Battelle)

**1:00 – 2:30 WEDNESDAY, APRIL 13<sup>TH</sup>**

**Environmental Laboratory Quality (ELQ) (Salons A-C) Chair: L. Doucet, EPA**

- Environmental Laboratory Quality Systems: Data Integrity Model and Systematic Procedures (R. DiRienzo, DataChem Laboratories, Inc.)
- The Interrelationship of Proficiency Testing, Interlaboratory Statistics and Lab QA Programs (T. Coyner, Analytical Products Group, Inc.)
- EPA FIFRA Laboratory Challenges and Solutions to Building a Quality System in Compliance with International Laboratory Quality Standard ISO 17025 (A. Ferdig, Mich. Dept. of Agriculture)

**Performance—Quality Systems Implementation (P-QSI) (Salon D) Chair: A. Belle, EPA**

- Implementing and Assessing Quality Systems for State, Tribal, and Local Agencies (K. Bolger, D. Johnson, L. Blume, EPA)

**1:00 – 2:30 WEDNESDAY, APRIL 13<sup>TH</sup> (continued)**

**Quality Initiatives in the EPA Office of Environmental Information (QI-OEI)** (Salons F-H) *Chair: J. Worthington, EPA*

- Next Generation Data Quality Automation in EPA Data Marts (P. Magrogan, Lockheed)
- The Design and Implementation of a Quality System for IT Products and Services (J. Scalera, EPA)
- Data Quality is in the Eyes of the Users: EPA's Locational Data Improvement Efforts (P. Garvey, EPA)

**A Win-Win-Win Partnership for Solving Environmental Problems (W3PSEP)** (Salon E) *Co-Chairs: W. Hunt, Jr. and K. Weems, NCSU*

- Overview of Environmental Statistics Courses at NCSU (B. Hunt, NCSU Statistics Dept.)
- Overview of the Environmental Statistics Program at Spelman College (N. Shah, Spelman)
- Student presentations: H. Ferguson and C. Smith of Spelman College; C. Pitts, B. Stines and J. White of NCSU

**Ambient Air Session III** (Sierra 5&6) *Chair: M. Papp, EPA*

- Trace Gas Monitoring for Support of the National Air Monitoring Strategy (D. Mikel, EPA)
- Comparison of the Proposed Versus Current Approach to Estimate Precision and Bias for Gaseous Automated Methods for the Ambient Air Monitoring Program (L. Camalier, EPA)
- Introduction to the IMPROVE Program's New Interactive Web-based Data Validation Tools (L. DeBell, Colorado State University)
- The Role of QA in Determination of Effects of Shipping Procedures for PM2.5 Speciation Filters (D. Crumpler, EPA)

**3:00 – 4:30 WEDNESDAY, APRIL 13<sup>TH</sup>**

**Topics in Environmental Data Operations (TEDO)** (Salons A-C) *Chair: M. Kantz, EPA*

- Ethics in Environmental Operations: It's More Than Just Lab Data (A. Rosecrance, Laboratory Data Consultants, Inc.)
- QA/QC of a Project Involving Cooperative Agreements, IAGs, Agency Staff and Contracts to Conduct the Research (A. Batterman, EPA)
- Dealing with Fishy Data: A Look at Quality Management for the Great Lakes Fish Monitoring Program (E. Murphy, EPA)

**Quality System Development (QSD)** (Salon D) *Chair: A. Belle, EPA*

- Development of a QA Program for the State of California (B. van Buuren, Van Buuren Consulting, LLC)
- Integrating EPA Quality System Requirements with Program Office Needs for a Practical Approach to Assuring Adequate Data Quality to Support Decision Making (K. Boynton, EPA)
- Introducing Quality System Changes in Large Established Organizations (H. Ferguson, EPA)

**Auditor Competence (AC)** (Salons F-H) *Chair: K. Orr, EPA*

- Determining the Competence of Auditors (G. Johnson, EPA)

**To Detect or Not Detect—What Is the Problem? (TDND)** (Salon E) *Chair: J. Warren, EPA*

- A Bayesian Approach to Measurement Detection Limits (B. Venner)
- The Problem of Statistical Analysis with Nondetects Present (D. Helsel, USGS)
- Handling Nondetects Using Survival Anal.(D. Helsel, USGS)
- Assessing the Risk associated with Mercury: Using ReVA's Webtool to Compare Data, Assumptions and Models (E. Smith, EPA)

**Ambient Air Session IV** (Sierra 5&6) *Chair: M. Papp, EPA*

- Status and Changes in EPA Infrastructure for Bias Traceability to NIST (M. Shanis, EPA)
- Using the TTP Laboratory at Sites with Higher Sample Flow Demands (A. Teitz, EPA )

**5:00 – 6:00 PM WEDNESDAY, APRIL 13<sup>TH</sup>**

**EPA SAS Users Group Meeting** Contact: Ann Pitchford, EPA

**8:30 – 10:00 THURSDAY, APRIL 14<sup>TH</sup>**

**Evaluating Environmental Data Quality (EEDQ)** (Salons A-C) *Chair: M. Kantz, EPA*

- QA Documentation to Support the Collection of Secondary Data (J. O'Donnell, Tetra Tech, Inc.)
- Staged Electronic Data Deliverable: Overview and Status (A. Mudambi, EPA)
- Automated Metadata Reports for Geo-Spatial Analyses (R. Booher, INDUS Corporation)

**Satellite Imagery QA (SI-QA)** (Salon D) *Chair: M. Cusanelli, EPA*

- Satellite Imagery QA Concerns (G. Brilis and R. Lunetta, EPA)

**Information Quality Perspectives (IQP)** (Salons F-H) *Chair: J. Worthington, EPA*

- A Body of Knowledge for Information and Data Quality (J. Worthington, L. Romero Cedeno, EPA)
- Information as an Environmental Technology – Approaching Quality from a Different Angle (K. Hull, Neptune and Co.)

**To Detect or Not Detect—What Is the Answer? (TDND)** (Salon E) *Chair: A. Pitchford, EPA, Co-Chair: W. Puckett, EPA*

- Using Small Area Analysis Statistics to Estimate Asthma Prevalence in Census Tracts from the National Health Interview Survey (T. Brody, EPA)
- Logistical Regression and QLIM Using SAS Software (J. Bander, SAS)
- Bayesian Estimation of the Mean in the Presence of Nondetects (A. Khago, University of Nevada)

**Ambient Air Workgroup Meeting** (Sierra 5&6) *Contact: Mike Papp, EPA*

NOTE: This is an all-day, closed meeting.

**10:30 – 12:00 THURSDAY, APRIL 14<sup>TH</sup>**

**Environmental Data Quality (EDQ)** (Salons A-C) *Chair: V. Holloman, EPA*

- Assessing Environmental Data Using External Calibration Procedures (Y. Yang, CSC)
- Groundwater Well Design Affects Data Representativeness: A Case Study on Organotins (E. Popek, Weston Solutions)

**Information Quality and Policy Frameworks (IQPF)** (Salons F-H) *Chair: L. Doucet, EPA*

- Modeling Quality Management System Practices to an Organization's Performance Measures (J. Worthington, L. Romero Cedeño, EPA)
- Development of a QAPP for Agency's Portal (K. Orr, EPA)
- Discussion of Drivers and Emerging Issues, Including IT, That May Result in Revisions to EPA's Quality Order and Manual (R. Shafer, EPA)

**Office of Water; Current Initiatives (OW)** (Salon D) *Chair: D. Sims, EPA*

- Whole Effluent Toxicity--The Role of QA in Litigation (M. Kelly, EPA, H. McCarty, CSC)
- Review of Data from Method Validation Studies: Ensuring Results Are Useful Without Putting the Cart Before the Horse (W. Telliard, EPA, H. McCarty, CSC)
- Detection and Quantitation Concepts: Where Are We Now? (Telliard, Kelly, and McCarty)

**Sampling Inside, Outside, and Under (SIOU)** (Salon E) *Chair: J. Warren, EPA*

- VSP Software: Designs and Data Analyses for Sampling – Contaminated Buildings (B. Pulsipher, J. Wilson, Pacific Northwest National Laboratory, R. O. Gilbert)
- Incorporating Statistical Analysis for Site Assessment into a Geographic Information System (D. Reichhardt, MSE Technology Applications, Inc.)
- The OPP's Pesticide Data Program Environmental Indicator Project (P. Villanueva, EPA)

**1:00 – 2:30 THURSDAY, APRIL 14<sup>TH</sup>**

**Information Management** (Salons A-C) *Chair: C. Thoma, EPA*

- Achieve Information Management Objectives by Building and Implementing a Data Quality Strategy (F. Dravis, Firstlogic)

**UFP Implementation** (Salon D) *Chair: D. Sims, EPA*

- Implementing the Products of the Intergovernmental DQ Task Force: The UFP QAPP (R. Runyon, M. Carter, EPA)
- Measuring Performance: The UFP QAPP Manual (M. Carter, EPA, C. Rastatter, VERSAR)

**Quality Systems Guidance and Training Developments (QSG)** (Salons F-H) *Chair: M. Kantz, EPA*

- A Sampling and Analysis Plan Guidance for Wetlands Projects (D. Taylor, EPA )
- My Top Ten List of Important Things I Do as an EPA QA and Records Manager (T. Hughes, EPA)
- I'm Here---I'm Free---Use Me! Use Me!—Secondary Use of Data in Your Quality System (M. Kantz, EPA)

**Innovative Environmental Analyses (IEA)** (Salon E) *Chair: M. Conomos, EPA*

- Evaluation of Replication Methods between NHANES 1999-2000 and NHANES 2001-2002 (H. Allender, EPA)
- Assessment of the Relative Importance of the CrEAM Model's Metrics (A. Lubin, L. Lehrman, and M. White, EPA)
- Statistical Evaluation Plans for Compliance Monitoring Programs (R. Ellgas, Shaw Environmental, Inc.; J. Shaw, EMCON/OWT, Inc.)

## Assessing Environmental Data Using External Calibration Procedure

Terry Smith

US EPA Analytical Operations Center, 1235 Jefferson Davis Highway, Arlington, VA 22202

Yan Yang

Computer Sciences Corporation, 15000 Conference Center Dr., Chantilly, VA 20151  
(703) 818-4524

*Under US Environmental Protection Agency (EPA) Contract Laboratory Program (CLP), laboratories are to analyze water, sediment, and soil samples from hazardous waste sites following the Contract Laboratory Program Statement of Work (CLP SOW). According to CLP methods, the analyses for polychlorinated biphenyls (PCBs) as Aroclors employ gas chromatography coupled with electron capture detector (GC/ECD) instrumentation for Aroclor qualitative identification and external standard calibration procedure for data quantitation. By using this method, CLP laboratories have successfully provided large amount of quality data to EPA throughout the years.*

*Often samples from various waste sites exhibit high Aroclor analyte peak area responses that exceed the upper limit of the calibration range established by the initial calibration. In order to provide quality data, the laboratories are to perform diluted re-analyses to bring the Aroclor analyte peaks within the calibration range. Occasionally laboratories use concentrations in place of the responses of the samples and the high standard of the initial calibration for comparison when considering sample dilutions. This can result in unnecessary dilutions of the samples and loss of time and resources at the laboratory.*

*This paper will present Aroclor analytical data provided by the CLP laboratory and examine the correlations between concentrations and area responses. By using Aroclor analyte peak area responses for sample dilution re-analysis determination, the laboratories can provide quality data to EPA, and at the same time, reduce analysis turn around time.*

### Introduction

US Environmental Protection Agency (EPA) Contract Laboratory Program (CLP) laboratories analyze water, sediment, and soil samples from hazardous waste sites following the Contract Laboratory Program Statement of Work (CLP SOW). The Statement of Work (SOW) is generated based on USEPA publication SW-846 that contains methods from sample preparation, extraction, extract cleanup techniques to sample target analyte qualitative identification and quantitation. According to CLP methods, the analyses for polychlorinated biphenyls (PCBs) as Aroclors employ gas chromatography coupled with electron capture detector (GC/ECD) instrumentation for Aroclor identification and external standard calibration procedure for data quantitation.

In the OLM04.2 SOW, analyses for Aroclors and individual pesticide analytes are combined into a single wide-bore capillary GC/EC analysis. For initial calibration, three concentration levels of pesticide reference standards are analyzed to determine pesticide retention time windows and calibration factors for quantitative analysis, and to establish detector linearity. Single point concentration standards for the major Aroclors are then analyzed to establish retention time windows and calibration factors. A set of three to five major individual Chlorinated Biphenyl isomers, representative of each Aroclor, is selected from each reference standard for retention time and calibration factor establishment.

When samples from certain hazardous waste sites exhibit high Aroclor analyte peak area responses that exceed the upper limit of the calibration range, the laboratories are to perform diluted re-analyses to bring the Aroclor analyte peaks within the calibration range. In the event of the single-point Aroclor initial calibration, samples are to be diluted when the area response of the largest peak in an Aroclor analyte is greater than the most intense single component pesticide analyte response in the initial calibration high point standard. For the Aroclor analytes that have three-point initial calibration, diluted re-analyses are required when any peak area response exceeds the corresponding peak area response of the high standard in the initial calibration.

### **Area Response vs. Concentration**

Occasionally laboratories use concentrations in place of the area responses of the samples and the high standard of the initial calibration for comparison when considering sample dilutions. In doing so, the laboratory chemist/analyst must have believed: a) The Aroclor analyte peak concentrations correlate well with the peak area responses of the sample; b) The upper limit of the calibration range for each Aroclor analyte peak is equivalent to 16 times of the adjusted Contract Required Quantitation Limit (CRQL) of the sample. In other words, when an Aroclor analyte peak concentration exceeds the upper limit of the calibration range, ie, the upper limit of the adjusted CRQL, the area response of this peak should exhibit greater response than either the most intense analyte peak in a single-point initial calibration or the corresponding peak of the high standard in a three-point initial calibration.

For most of the Aroclor analyses, the correlations between the peak concentrations and area responses are such that the higher the Aroclor analyte peak concentration, the greater the area response. However, there are situations where such correlations are invalid that the peak concentration and peak area response exhibit values opposite to each other. It is easy to understand when we look at the following concentration equation used for soil/sediment sample

$$\text{Concentration} \sim \mu \text{g/Kg} \sim (\text{Dry weight basis})$$

$$\sim \frac{(A_x)(V_t)(Df)(GPC)}{(CF)(V_i)(W_s)(D)}$$

quantitation.

It is noted that concentration is proportional to Aroclor analyte peak area, **A<sub>x</sub>**; and inversely proportional to its calibration factor, **CF**, from the single-point initial calibration or the mean calibration factor from the three-point initial calibration for the corresponding peak. Thus, it is possible for a resultant Aroclor peak concentration to be high, and at the same time, its area

response to be low so that it is less than that of the upper limit of the calibration range. According to the SOW, the contract CRQL is proportional to the concentration of the low standard in initial calibration. And the adjusted CRQL for the soil/sediment sample can be calculated from the following SOW equation:

$$\text{func}\{\text{stack}\{\text{Adjusted \# CRQL}\} \sim \sim \text{stack}\{\text{Contract \# CRQL}\} \\ \sim \text{times} \sim \{(W_x) \cdot (V_t) \cdot (V_y) \cdot (Df)\} \\ \text{over} \{(W_s) \cdot (V_c) \cdot (V_i) \cdot (D)\}\}$$

The initial calibration standards are prepared at three concentration levels such that the highest standard concentration is 16 times of the low standard concentration. Since the CRQL is proportional to the concentration of the standard, the upper limit of the CRQL can be determined to be 16 times of the CRQL. The same multiplier can be extended to the adjusted CRQL which is directly proportional to the CRQL. However, for three-five Aroclor peaks with various peak intensities, i.e. area responses, used for quantitation, it is not practical to precisely determine the multiplier to be used for calculating the upper limit of the calibration range for each peak. Therefore, to evaluate the diluted re-analysis situation by comparing either peak or the mean analyte concentration in a sample with the upper limit of the adjusted CRQL is scientifically inappropriate.

## Laboratory Data

Soil samples from certain waste sites are analyzed at two of the CLP laboratories, Lab A and Lab B, following the analytical procedures outlined in OLM04.2 SOW. 30 g aliquot of sample is spiked with the surrogates, tetrachloro-m-xylene (TCX) and decachlorobiphenyl (DCB); and then mixed with anhydrous sodium sulfate, extracted with a 1:1 (v/v) acetone/methylene chloride solvent mixture by sonication extraction method. The extract is filtered, concentrated, and cleaned up by GPC. The extract is then solvent exchanged into hexane, and adjusted to a final volume of 1.0 mL or 0.25 ml for the modified analysis. 1 ul of the sample extract is injected onto the GC with the instrument conditions listed in Table A. Both laboratories use the GC columns manufactured from Restek with the specifications listed in Table B.

Data system interfaced to the GC/EC provided chromatograms and data integrations reports. Figure 1 shows the chromatogram of GC column 1 for Lab A sample1. Three Aroclor-1242 analyte peaks are identified. Figure 2 shows the chromatogram of the primary GC column for Lab B sample 2. Three Aroclor analytes, Aroclor-1242, Aroclor-1248 and Aroclor-1254, are identified with three peaks, four peaks and four peaks for the respective analyte. The external calibration standards at the following specified concentration levels were injected onto the GC: single point Aroclor standards are analyzed at 0.1 ng/uL where three point Aroclor standards with modified analysis are at 0.1, 0.4 and 1.6 ng/uL respectively. For Methoxychlor, the most intense peak in the individual standard mixture, the three concentration levels are 0.05, 0.2 and 0.8 ng/uL. The Aroclor analyte calibration factor or mean calibration factors and sample



Aroclor analyte concentrations are calculated using the equations in SOW. The following tables show the Aroclor analyte quantitation summary for the samples from both laboratories. Table I & II are for samples with single-point initial calibration while Table III is for the sample with three-point initial calibration for Aroclor-1242 and Aroclor-1254 and single-point initial calibration for Aroclor-1248. Methoxychlor high standard CS3 area responses for both GC columns are tabulated along with the upper limits of the adjusted CRQLs which are determined using adjusted CRQLs multiplying by 16. Figure 3 shows Aroclor-1248 area response, concentration and calibration Factor comparison for Lab B sample 2 on column 1. For comparison purpose, the concentration values on the graph are 100 times of the true values. Figure 4 and Figure 5 are Aroclor-1242 analyte peak area response comparisons for Lab A samples 1 and 2. Figure 6 and Figure 7 are Aroclor-1248 analyte peak area response comparisons for Lab B samples 1 and 2. Figure 8 and Figure 9 are Aroclor-1242, Aroclor-1254 analyte peak area response comparisons for Lab B sample 2.

## Discussion

Figure 3 is an illustration of the relationships between area response and concentration for Lab B sample 2. It is noted that the area response of peak 3 is relatively greater than that of peak 4, however, the concentration of peak 3 is much lower than that of peak 4 due to its calibration factor being less. For peak 1 and 2, the area responses correlate well with their concentrations due to their similar values in calibration factors. Again the correlation between the area response and concentration shown in the graph has demonstrated that concentration of an analyte peak can be high when its area response is low and vice versa.

Table I and II are the summaries for samples from the same Region. Sample 1 has been diluted once with dilution factor of 5.0. It is noticed that the mean concentrations of Aroclor-1242 analyte for both columns are greater than the upper limit of the adjusted CRQL. On the other hand, peak 3, the largest peak for column 1, has peak area response of  $2.65\text{E}+08$  which is less than  $2.96\text{E}+08$ , the area response of Methoxychlor, the most intense peak in the high standard CS3 for column 1. The column 2 peak area responses have shown the same trend. Figure 4 shows clearly that all Aroclor-1242 analyte peak area responses are less than that of the Methoxychlor in CS3 for both columns. Therefore, no further dilution is required for this sample based on the peak area response comparison. Yet the lab performed the diluted re-analysis based on the concentration comparison.

For sample 2, the mean concentrations of Aroclor-1242 analyte for both columns are greater than the upper limit of the adjusted CRQL as shown in Table I. At the same time, peak 3, the largest peak for column 1, has peak area response of  $5.85\text{E}+08$  which is greater than  $2.96\text{E}+08$ , the area response of Methoxychlor, the most intense peak in the high standard CS3 for column 1. For column 2, none of the peaks have area responses exceeded that of Methoxychlor. Figure 5 shows clearly that peak 3 for column 1 has higher area response than Methoxychlor. Therefore, diluted re-analysis for this sample is required regardless of the concentration comparison. Sample 1 from Lab B has shown higher mean concentrations of Aroclor-1248 for both GC columns in Table II. However, none of the peaks have area responses exceed that of Methoxychlor in CS3 for both GC columns as shown in Figure 6. Therefore, diluted re-analysis

is not required for sample 1 at Lab B. As a result, Lab B could have saved time and resources if sample 1 was not diluted for analyte Aroclor-1248.

Sample 2 from Lab B was analyzed with modified analysis that two Aroclor analytes, Aroclor-1242 and Aroclor-1254, are quantified using mean calibration factors from the three-point initial calibration. Yet Aroclor analyte Aroclor-1248 was quantified using its calibration factor from one-point initial calibration. Therefore, for analyte Aroclor-1248, the comparison of the area responses with Methoxychlor shown in Figure 7 manifests that none peak exceeds the upper limit of the calibration range even if the concentration comparison demonstrates the opposite. For analytes Aroclor-1242 and Aroclor-1254, all the mean concentrations in Table III show that they are actually less than the upper limit of the adjusted CRQLs. However, peak 1 of analyte Aroclor-1242, the largest peak for column 1, has greater area response than its high standard CS3. The same has occurred for peak 1 on column 2. The comparisons for both columns are clearly shown in Figure 8. For analyte Aroclor-1254, it is peak 3 for column 2 whose area response edges out that of the high standard CS3 as shown in Figure 9. Therefore, a diluted re-analysis is sought for this sample at Lab B.

## **Conclusion**

The Laboratory data have illustrated the correlations between Aroclor analyte peak response and concentration that higher Aroclor analyte peak concentration does not translate into a greater analyte peak area response. Neither the individual peak concentration nor the mean analyte concentration can be viable for evaluating sample dilution situation. The Laboratory data have also demonstrated that using concentrations in place of the responses of the samples and the high standard of the initial calibration for comparison can result in unnecessary sample dilutions and consequently loss of time and resources at the laboratory. Therefore, the laboratories should always compare peak area responses when considering sample dilutions. By doing so, the laboratories can provide quality data to EPA, and at the same time, reduce analysis turn around time.

## **Reference**

1. United States Environmental Protection Agency, "Statement of Work for Organics Analysis", Multi-Media, Multi-Concentration, OLM04.2 Exhibit D Pesticide/Aroclor, May 1999.

## Figures & Tables



Figure 1 Chromatograms for Lab A sample1 on GC column 1



Figure 2 Chromatogram for Lab B sample 2 on GC column 1

**Table A GC Instrument Condition**

Carrier gas (He)	5 mL/min.
Injector temperature	205EC
Detector temperature	290EC
Initial temperature	150EC, hold 0.5 min.
Temperature program	150EC to 275EC at 5EC/min.
Final temperature	275EC, hold 10 min.

**Table B GC Column Specification**

Column Dimensions	Column 1	Column 2
Length (m)	30	30
ID (mm)	0.53	0.53
Film Thickness (um)	0.5	0.42

**Table I Lab A Sample Aroclor Quantitation Summary**

Analyte	Peak	RT (min)	Area Response	Concentration (ug/kg)	Mean Concentration	Methoxychlor CS3 Area Response	Upper limit Adjusted CRQL (ug/kg)
Sample 1	1	6.34	1.75E+07	940			
Aroclor-1242	2	7.27	2.42E+08	6600	3900	2.96E+08	3400
Column 1	3	8.42	2.65E+08	4200			
	1	9.52	3.66E+08	7300			
Column 2	2	10.68	4.31E+08	3600	4300	9.18E+08	3400
	3	11.02	1.11E+08	1900			
Sample 2	1	6.34	1.25E+08	1500			
Aroclor-1242	2	7.27	3.74E+08	2500	2200	2.96E+08	660
Column 1	3	8.42	5.85E+08	2500			
	1	9.52	6.12E+08	2400			
Column 2	2	10.68	8.73E+08	1800	1700	9.18E+08	660
	3	11.02	1.68E+08	750			

**Table II Lab B Sample Aroclor Quantitation Summary**

Analyte	Peak	RT (min)	Area Response	Concentration (ug/kg)	Mean Concentration	Methoxychlor CS3 Area Response	Upper limit Adjusted CRQL (ug/kg)
Sample 1	1	7.32	4.49E+04	820			
Aroclor-1248	2	7.84	4.76E+04	870			
Column 1	3	8.89	4.22E+04	670	790	2.00E+05	740
	4	9.43	3.84E+04	720			
	5	10.42	2.96E+04	880			
	1	9.98	2.79E+04	1000			
Column 2	2	10.3	3.53E+04	1100	1100	1.36E+05	740
	4	12.79	2.49E+04	1100			
	5	13.4	2.19E+04	1200			

**Table III Lab B Sample Aroclor Quantitation Summary**

Analyte	Peak	RT (min)	Area Response	Concentration (ug/kg)	Mean Concentration (ug/kg)	Methoxychlor or Aroclor CS3 Area Response	Upper Limit Adjusted CRQL (ug/kg)
Sample 2	1	7.19	4.95E+04	440			
Aroclor-1248	2	7.71	3.35E+04	300	360	1.96E+05	340
Column 1	3	8.76	4.15E+04	320			
	4	9.3	4.09E+04	380			
	1	10.02	1.18E+04	320	510	6.95E+04	340
Column 2	2	10.34	1.91E+04	460			
	3	11.13	3.93E+04	760			
	1	4.52	1.54E+04	390		1.04E+04	
Aroclor-1242	2	5.33	1.07E+04	170	240	1.71E+04	340
Column 1	4	6.71	9.63E+03	150		1.73E+04	
	1	6.18	6.46E+03	400		4.47E+03	
Column 2	2	7.26	5.86E+03	220	270	7.06E+03	340
	4	8.78	5.03E+03	190		7.15E+03	
	2	10.28	3.01E+04	180		4.64E+04	
Aroclor-1254	3	10.99	4.11E+04	200	160	6.12E+04	340
Column 1	4	11.47	2.27E+04	130		4.98E+04	
	5	12.16	3.50E+04	140		7.01E+04	
	2	12.83	1.65E+04	270		1.68E+04	
Column 2	3	13.43	2.02E+04	290	220	1.92E+04	340
	4	13.92	8.15E+03	180		1.24E+04	
	5	14.82	1.31E+04	150		2.38E+04	

## **Groundwater Well Design Affects Data Representativeness—A Case Study on Organotins**

**Emma P. Popek, Ph.D. and Tracy L. Walker, R.G.**

Representativeness is a measure of the degree to which data accurately and precisely represent a characteristic of a population, a parameter variation at sampling point, a process condition, or an environmental condition. Poorly understood material properties may bring about environmental system designs that preclude the collection of representative data. The quality of the environment system design thus affects the quality of collected data, which, while believed to be accurate, precise and representative, in fact, may not correctly represent the true environmental condition. If such design error is recognized, the data, although valid, could be unusable for project decisions. However, if the design error is never discovered and acknowledged, the use of unrepresentative data may cause wrong decisions related to site investigation, monitoring, and remediation. The monetary costs of such errors can be significant.

A rare case study supports our statement. A landfill at a former Naval shipyard in California is undergoing remediation and closure. Among the landfill groundwater contaminants monitored since 1992 are organotins, compounds used in ship antifouling paints. Several organotin compounds were detected in monitoring wells after they were installed. Further monitoring did not find any presence of organotins in the groundwater at the site until recently when new wells were installed, sampled, and low levels of organotin compounds detected. Research has shown that the source of organotins in groundwater is the polyvinyl chloride casings of the monitoring wells; the antifouling paint may be ruled out as a source.

The unnoticed fact that organotin data were not representative of the true groundwater conditions at the landfill caused unnecessary monitoring conducted at a significant cost. Moreover, organotins became constituents of concern under California regulations and may require monitoring over the next 30 years. The use of stainless steel casings for groundwater monitoring wells would have been a better environmental monitoring system design choice, averting the collection of unrepresentative data.

## **Modeling Quality Management System Practices to an Organization's Performance Measures**

*Jeffrey Worthington – OEI Director of Quality, Office of Planning, Resources, and Outreach, Office of Environmental Information, US Environmental Protection Agency*

*Lorena Romero Cedeño – Program Analyst, Office of Planning, Resources, and Outreach, Office of Environmental Information, US Environmental Protection Agency*

### **Abstract:**

Management systems in the public (Federal, State, etc.) and private sectors often include multiple planning documents replete with objectives, goals, and criteria. When they are used, performance measures generally involve a discrete reporting system and a variety of required, quantitative and verifiable targets and accountabilities. Similarly, when quality planning is used, it also includes setting goals, whether the goals are for systems, programs, projects or some other relevant component, and measurement and assessment criteria for each of the selected goals. Aligning the management objectives and measures with quality objectives and measures can strengthen the overall management processes for an organization and its resources.

This presentation reviews the various management and quality systems and some of the associated accountabilities. The rationales for various objectives and measures and the challenges to aligning them will also be reviewed. It is important to recognize that a quality system and its associated objectives represent only one of many management systems. As such, the quality system is unlikely to address all of an organization's performance measurement needs. This presentation provides a method for reviewing an organization's performance and quality measures, and aligning them.

## **Development of a Quality Assurance Project Plan for the Agency's Portal**

*Kimberlie Orr, Environmental Protection Specialist, Quality Staff, Office of Environmental Information, US Environmental Protection Agency*

### **Abstract:**

Over the past few months, the Office of Environmental Information (OEI) Quality Staff has been working with the OEI Program Management Office support contractor to discuss the development of a Quality Assurance Project Plan (QAPP) for the Environmental Information Integration and Portal Development Project. The QAPP outline we are developing incorporates the lifecycle stages of a project, from the determination of objectives to information on agency reviews to the steps needed to prepare for dissemination. Putting together a draft for such a multi-faceted project has been both illuminating and challenging. During this presentation, we will discuss how the draft QAPP took shape.



## **Discussion of drivers and emerging issues, including IT, that may result in revisions to EPA's Quality Order and Manual**

*Ron Shafer, Operations Research Analyst, Quality Staff, Office of Environmental Information, US Environmental Protection Agency*

### **Abstract:**

The development of a National Exchange Network will provide EPA, its partners, and the public with better access to Agency data and information. OMB's Data Quality Act (and EPA's resulting Information Quality Guidelines) gives stakeholders and the public the ability to dispute EPA's disseminated information. During this presentation, we will discuss the results of evaluations of the effectiveness of EPA's Quality Program to assure quality of data, databases and information products being disseminated to the public, and will discuss emerging issues. Suggestions will be provided to revise Order 5360.1 to include the recent Information Quality Guidelines, and to develop an Intranet Web site that provides quality assurance support tools and examples.

## **Abstract**

### **Whole Effluent Toxicity - The Role of Quality Assurance in Litigation**

Marion Kelly, U.S. Environmental Protection Agency  
Harry B. McCarty, CSC Environmental Programs Group

On October 16, 1995, the EPA Office of Water (OW) amended the “Guidelines Establishing Test Procedures for the Analysis of Pollutants” (40 CFR Part 136) to add a series of standardized whole effluent toxicity (WET) test methods to the list of Agency approved methods for conducting required testing of aqueous samples under the Clean Water Act. The rule was challenged in court by a coalition of industrial dischargers. To resolve the litigation, EPA entered into settlement agreements with the various parties and agreed to publish a method guidance document and a variability guidance document, conduct an interlaboratory variability study, publish a peer-reviewed interlaboratory variability study report, address pathogen contamination, propose specific technical method changes, and propose to ratify or withdraw WET test methods evaluated in the interlaboratory variability study. The objectives of the interlaboratory study were to generate data to:

- Characterize the interlaboratory variability of the 12 WET methods targeted in the study
- Estimate the rate at which participating laboratories successfully completed WET tests initiated, and
- Estimate the rate at which WET tests indicate “toxicity” is present when measuring non-toxic samples.

EPA completed the study and promulgated revised WET methods in a final rulemaking in November 2002, which was challenged by the same parties in October 2004. The U.S. Court of Appeals for the District of Columbia Circuit ruled in favor of EPA on December 15, 2004.

A large measure of EPA’s success in this case is due to the Agency’s careful attention to the four components of Deming’s “quality circle” of *plan, do, check, and act*. In carrying out method development and validation activities for WET methods, the Office of Water was careful to plan and document decisions made during each step of the process. OW developed six data quality objectives for the study. OW took the added steps to subject the study plan for the interlaboratory study to a formal peer review, as well as a review by the litigants set forth in the settlement agreement.

In total, 12 WET methods were tested with 56 laboratories evaluating an average of 3 methods each, through the analysis of more than 700 samples. Over 300,000 individual biological, physical, and chemical measurements were conducted and reported in the study. All data from the interlaboratory study were subjected to formal validation procedures, many custom-designed for this study, to determine if the results met the data quality objectives identified in the study plan. The validation effort included comparison of results calculated in the lab and by EPA. While inconsistencies in result calculations were observed in 54% of tests, the majority of those inconsistencies were minor (63% resulting in <1% difference). All the inconsistencies were identified and corrected, resulting in a robust data set. Those data were

tabulated in a study report that was subjected to peer review before the final rulemaking.

The final data set and the supporting quality assurance effort were used by EPA and attorneys from the U.S. Department of Justice to defend EPA's 2002 promulgation of the WET methods. The Court took note of EPA's efforts to validate the results and conduct peer reviews in its December 2004 decision in favor of the Agency.

## **Review of Data from Method Validation Studies: Ensuring Results Are Useful Without Putting the Cart Before the Horse**

William Telliard, U.S. EPA, Engineering and Analysis Division  
Harry B. McCarty, CSC Environmental Programs Group

The EPA Office of Water (OW) approves analytical methods for use in monitoring compliance with Clean Water Act programs. When new regulatory needs are identified, OW works with various organizations to develop and validate methods for new contaminants, methods with greater sensitivity, or methods for new matrices. Method validation typically involves one or more single-laboratory studies designed as a proof of the analytical concept, followed by an interlaboratory study designed to test the ruggedness of the method and to produce data that can be used to develop or confirm quality control (QC) acceptance criteria for various aspects of the method.

In order to withstand the rigors of proposal and public comment, OW subjects all data from interlaboratory studies to extensive review procedures designed to ensure that the results meet the objectives of the study. However, those procedures usually are based on review procedures used for routine analyses of samples by established methods and they have to be adjusted for the fact that the purpose of the study is to develop data from which to derive acceptance criteria, rather than checking the data against existing criteria. Otherwise, you would be checking data against criteria that do not yet exist, or putting the cart before the horse.

EPA and the voluntary consensus standards bodies (VCSBs) that develop and publish analytical methods often conduct Interlaboratory studies using volunteer participants, especially for methods involving expensive instrumentation. A case in point is EPA's recent interlaboratory study of Method 1668A, a high resolution GC/MS procedure for the analysis of 209 polychlorinated biphenyl (PCB) congeners.

The use of volunteer laboratories further complicates the interlaboratory validation process. By virtue of volunteering to participate, a laboratory has no financial incentive to complete the study in a given time frame, to provide results in the format requested for the study, or to respond to inquiries after the data are delivered. During an international validation study of EPA Method 1613 (dioxins and furans) in 1990, it took 11 months to obtain results from 20 of the 21 laboratories that volunteered to participate. One laboratory never provided results at all. In the case of the study of EPA Method 1668A, 3 of 14 laboratories still have not submitted results over two years after the study began.

Other problems common to interlaboratory validation studies include limited experience with the exact methodology by some of the laboratories, inexperience with difficult sample matrices of interest to EPA (e.g., tissues or biosolids), and failure to describe how the samples were analyzed in sufficient detail to allow reviewers to identify potential weaknesses in the techniques or the written method.

OW has used various approaches to overcome the difficulties in validation studies, including careful planning and explicit documentation of the study design and expectations,

prequalification questionnaires for potential volunteers, offsetting participation costs by providing study-specific standards, holding conference calls with participants prior to the start of the study, and frequent communications with the participants during the study and during the review process. OW has considered videotaping important aspects of some procedures to ensure consistency among participants.

## **Detection and Quantitation Concepts Where Are We Now?**

William Telliard and Marion Kelly, U.S. EPA, Engineering and Analysis Division  
Harry B. McCarty, CSC Environmental Programs Group

In response to the 1999 promulgation of EPA Method 1631 for the analysis of mercury at low levels, an industry group filed a lawsuit challenging promulgation of the method as well as the general procedures used by EPA to establish detection and quantitation limits. EPA entered into a settlement agreement that required the Agency to assess its concepts and procedures for determining the detection and quantitation limits in Clean Water Act Programs. The results of that assessment were published in March 2003 along with a proposal to revise the existing procedures for determining detection and quantitation limits.

In November 2004 EPA withdrew the proposal to revise the definition of the method detection limit (MDL) and formalize a definition of the minimum level of quantitation (ML). EPA concluded that approaches other than those set forth in the 2003 proposal have the potential for addressing concerns regarding development and use of detection and quantitation limits, and that those approaches warrant further consideration and refinement.

Despite the ongoing efforts of voluntary consensus standards bodies (VCSB) and based on public comments, there is no agreement among critics of the existing MDL and ML procedures about what changes should be adopted by the Agency for use in Clean Water Act programs. Therefore, EPA agreed to continue to work collaboratively with stakeholders on these issues. This collaborative process began in the Fall of 2004, when a neutral third party was hired to conduct a situation assessment to assess the feasibility and develop a stakeholder consultative process on detection and quantitation. Thirty-seven stakeholders from industry, state, publicly owned treatment works, environmental laboratories, environmental groups, and others were interviewed. Stakeholders recommended that the Agency establish a Federal Advisory Committee (FAC) to address the calculation and use of detection and quantitation limits under the Clean Water Act.

The Situation Assessment and the Office of General Council's Alternative Dispute Resolution Law Office recommended that a FAC be established with approximately 20 members with a balanced representation of views. The FAC will include qualified, senior-level professionals from state government; environmental professionals; regulated industry; environmental laboratories; POTWs; and the environmental community. There will be a technical workgroup that will make recommendations to the FAC. Neutral technical experts will help the committee understand technical concepts and provide assistance, but will not be members of the committee or participate in the committee's deliberations.

On December 29, 2004, EPA invited nominations of qualified candidates to be considered for appointment to the Federal Advisory Committee on Detection and Quantitation Approaches and Uses in Clean Water Act Programs. EPA conducted a half-day public meeting on January 26, 2005, to present the findings and recommendations in the Situation Assessment Report on Detection and Quantitation Approaches and Uses in Clean Water Act Programs. The

first meeting of the FAC is projected for late Spring or early Summer 2005, in Washington, DC.

## **Visual Sample Plan (VSP) Software: Designs and Data Analyses for Sampling Contaminated Buildings**

Brent A. Pulsipher, John E. Wilson, Richard O. Gilbert, Lisa Nuffer, Nancy Hassig

Statistical Sciences  
Pacific Northwest National Laboratory  
Richland, WA

### **Abstract**

A new module of the Visual Sample Plan (VSP) software has been developed to provide sampling designs and data analyses for potentially contaminated buildings. An important application is assessing levels of contamination in buildings after a terrorist attack. This new module, funded by DHS through the Combating Terrorism Technology Support Office, Technical Support Working Group, was developed to provide a tailored, user-friendly and visually-orientated buildings module within the existing VSP software toolkit, the latest version of which can be downloaded from <http://dgo.pnl.gov/vsp>. In case of, or when planning against, a chemical, biological, or radionuclide release within a building, the VSP module can be used to quickly and easily develop and visualize technically defensible sampling schemes for walls, floors, ceilings, and other surfaces to statistically determine if contamination is present, its magnitude and extent throughout the building and if decontamination has been effective.

This paper demonstrates the features of this new VSP buildings module, which include: the ability to import building floor plans or to easily draw, manipulate, and view rooms in several ways; being able to insert doors, windows and annotations into a room; 3-D graphic room views with surfaces labeled and floor plans that show building zones that have separate air handling units. The paper will also discuss the statistical design and data analysis options available in the buildings module. Design objectives supported include comparing an average to a threshold when the data distribution is normal or unknown, and comparing measurements to a threshold to detect hotspots or to insure most of the area is uncontaminated when the data distribution is normal or unknown.



## **1.0 Introduction**

With increasing risk of a biological, chemical, or radiological terrorist event, readiness efforts are underway throughout the World. The ability to quickly identify the contaminant of concern, determine extent of contamination, decontaminate an area, and evaluate the effectiveness of decontamination has a significant effect on health risks and economical consequences. Sampling will play a major role in obtaining sufficient information to make informed confident decisions throughout the entire response and restoration cycle. Guidelines, methods, and tools are needed to support all data gathering and analysis efforts.

The US Environmental Protection Agency (EPA) and the Department of Homeland Security (DHS) recognize the need for such guidelines, methods, and tools for a variety of chemical, biological, and radiological release scenarios. Through the Technical Support Working Group (TSWG), they are supporting the development of specific methods and tools aimed at facilitating sampling designs within buildings. These methods are being incorporated within Visual Sample Plan (VSP), a DQO-based statistical sampling design software package sponsored by the Department of Energy, Department of Defense, EPA, and DHS for developing defensible sampling schemes for soils, surface water, surfaces, and sites suspected of containing unexploded ordnance.

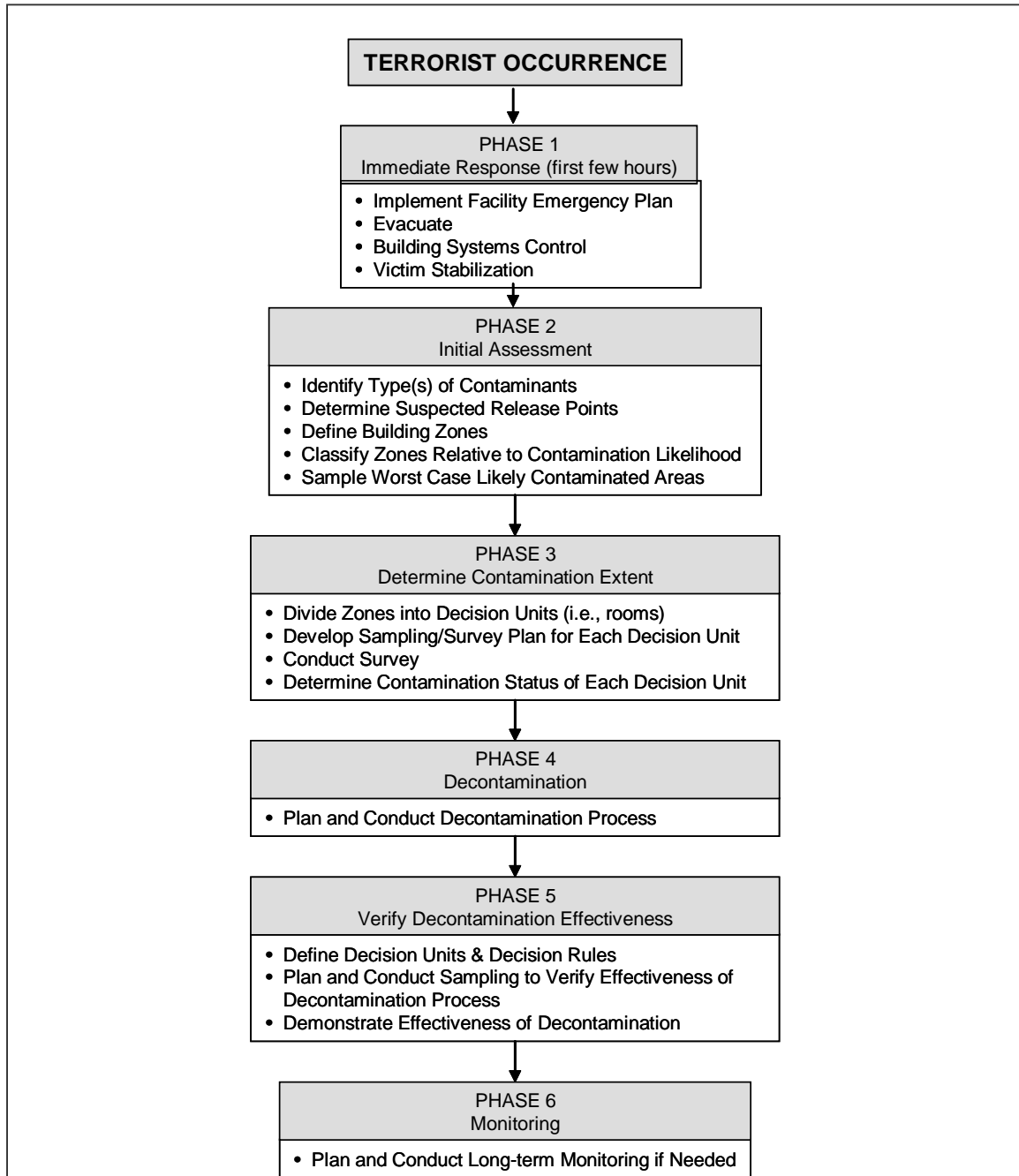
## **2.0 Sampling Requirements In Response to a Terrorist Chem/Bio/Rad Event**

When planning for or responding to a contaminant spreading event, sampling may be employed throughout several phases of the characterization and restoration cycle. It is important to quickly and easily develop and visualize technically defensible sampling schemes to identify and determine magnitude and extent of contamination to guide initial response, decontamination, and verification of decontamination effectiveness. Figure 1 shows several phases of the response and restoration process from a sampling design perspective.

During Phase 1, the immediate response phase, no sampling is envisioned. Only victim stabilization, building controls, initiation of the emergency response plans are considered. During Phase 2, Initial Assessment, some judgment sampling will probably be performed to identify the contaminant(s) of concern and get an initial estimate of the magnitude of contamination. This judgment sampling approach will lead to samples of any visible suspect material and in most likely contaminated areas. A statistically representative set of samples is probably not a viable nor desirable approach during this initial assessment phase.

In Phase 3, the primary objective of sampling is to determine the extent of contamination. A more systematic, well planned approach to sampling is needed in this phase to guide future decontamination decisions. Both judgmental and statistically-based sampling schemes may be appropriate during this phase of the response depending on the decision rules employed.

In Phase 4, Decontamination, the primary sampling may be associated with the placement of spiked samples that are evaluated to guide the decontamination process. For example, in the case of an anthrax release, one decontamination technology may involve room fumigation. Samples spiked with spore colonies may be strategically placed within a room prior to fumigation and examined after fumigation to determine whether the decontamination procedure effectively destroyed the bio-pathogens or additional fumigation is required.



**Figure 1. Phased Response to Chem/Bio/Rad Event – Sampling Perspective.**

During Phase 5, the primary objective of sampling is to verify the effectiveness of the decontamination process such that no contaminants exist above threshold levels of concern. Statistical sampling design approaches will definitely be appropriate during this phase. Finally, in Phase 6, sampling schemes may need to be derived to monitor for contaminants that may have escaped detection and eradication.

### **3.0 Decision Rules and Desired Statistical Confidence Statements**

The Data Quality Objectives (DQO) process helps identify what decisions must be supported by data during each phase of the response and restoration cycle. By following a DQO approach, specific decision rules can be developed and required confidence levels can be defined. As illustrated above, there are a variety of decisions and sampling objectives for the various phases of the response and restoration cycle. An optimal sampling scheme for identifying the contaminants would not necessarily be optimal for determining the extent of contamination. Each sampling event usually supports a different decision and can employ a different decision rule with a specific required level of confidence and tolerance for error.

Decision rules specify

- the contaminant of concern,
- some parameter of interest (i.e., mean, median, maximum),
- the action level (regulatory limit),
- and alternative actions.

In the case of a terrorist bio/chem/rad event, the parameters of interest will most likely be the maximum, mean, or some percentile. Without clear regulatory guidance, one way to develop the decision rule is to determine what statement one would like to be able to make after the data have been analyzed. For example, do you want to state that you are 95% confident that the true contaminant mean is less than the regulatory limit? Or perhaps you want to state that you are 95% confident that at least 99% of the surface area is uncontaminated (less than some limit). The statistical sampling requirements will entirely depend on the decision rule and the type of desired confidence statement. Decision rules based on the mean or on maximums and percentiles are discussed below.

#### **3.1 Mean-Based Decision Rules and Confidence Intervals**

When the sampling objective is primarily to gather sufficient data to estimate the mean, the three most common decision rules and estimation objectives are

- Comparing the mean against some predefined regulatory threshold,
- Comparing the mean against some reference (background) mean, or
- Bounding the estimate of the mean by some upper and/or lower limit.

Table 1 identifies the decision criteria, the sampling design criteria, and the resulting decision rules and confidence statements that can be made for each of these mean-based decision rules. In practice, the decision makers and stakeholders must decide which confidence statements and decision rules they want to apply before an optimal sampling design can be derived. A number of statistical sampling designs could be applicable

**Table 1. Possible Decision Rules and Resulting Confidence Statements**

Decision Criteria	Design Criteria	Resulting Decision Rule and Confidence Statement
<b>Decision Rules Based on Means</b>		
<b>Mean vs. Threshold</b>	Determine n such that there is no more than a 100P% chance of deciding the true mean is < limit if the true mean is > limit; and only a 100Q% chance of deciding the true mean is > limit if it is significantly <limit.	If we are 100(1-P)% confident that the mean is < limit, then the decision unit is deemed to be uncontaminated; otherwise it is deemed to be contaminated.
<b>Mean vs. Background Mean</b>	Determine n such that there is no more than a 100P% chance of deciding the true mean = the background mean when the true mean is significantly greater than background; and only a 100Q% chance of deciding the mean is > background when there is no significant difference.	If we are 100(1-P)% confident that the mean is > the background mean, then the decision unit is contaminated; otherwise it is deemed to be uncontaminated.
<b>Mean: Confidence Interval or limit</b>	Determine n such that the halfwidth of a 100(1-P)% confidence interval is no greater than X, e.g. estimate the mean to within $\pm X$ with 100(1-P)% confidence.	One is 100(1-P)% confident that the true mean is within the upper and lower limit of the confidence interval.
<b>Decision Rules Based on Individual Sample Results, Maximums, or Percentiles</b>		
<b>Individual Sample Results: Hotspot</b> Decision Criteria	Develop Sampling Grid & n to ensure 100(1-P)% confident of obtaining at least 1 sample from a hotspot of a specified size and shape if one exists.	If all measured values < limit, one can be 100(1-P)% confident that a hotspot of specified size/shape does not exist. If any measurement > limit, the decision unit is deemed to be contaminated; hotspot(s) exists.
<b>Percentile (UTL) or Maximum:</b> Less than some proportion of area contaminated	Determine n such that if all of the n measurements are < limit, then there is < 100P% chance that > R% of the possible measurements are > limit. May use Nonparametric UTL or Parametric UTL for finite or infinite populations; also compliance sampling approach.	If the maximum measured value (for nonparametric UTL) or the computed UTL is < limit, then one is 100(1-P)% confident that no more than R% of all possible measurements is > limit. Otherwise, the decision unit is deemed to be contaminated.
<b>Individual Sample Results:</b> No exceedances allowed.	Determine n such that if all of the n measurements are < limit, there is <100 P% chance of any of the remaining N-n possible measurements to be > limit.	If all n results are < limit, then one is 100(1-P)% confident that ALL remaining unsampled areas are < limit. Otherwise the decision unit is deemed to be contaminated.
<b>Individual Sample Results:</b> Some sample exceedances allowed.	Determine n and c such that if the true number of exceedances in the population are < X, then there is <100P% chance of getting more than c exceedances; and only a 100Q% chance of getting $\leq c$ sample exceedances if the number of exceedances in the population is > Y.	If there are c or fewer sample results that are > limit, then there is < 100P% chance that there are more than X exceedances in the population so the contamination is deemed to be acceptable. Otherwise the decision unit is deemed to be contaminated.

depending on the assumptions, constraints, and sampling technologies. These designs might include random sampling, systematic grid sampling, collaborative sampling, stratified sampling, or sequential sampling.

### **3.2 Decision Rules Based on Individual Sample Results, Maximums, or Percentiles**

In some cases, the primary sampling objective is to estimate the maximum contamination level or to compare the maximum or a percentile estimate (Upper Tolerance Limit, UTL) against some threshold of concern. Common decision rules involve

- Comparing individual measurements against some predefined regulatory threshold to identify potential contaminant hotspots,
- Comparing the maximum value against a threshold, or
- Comparing an estimated upper confidence limit on some percentile against some threshold.

Table 1 also shows the decision and sampling design criteria and the resulting decision rules and confidence statements for each of these objectives. There are a number of sampling design methods that can be used for some of these decision rules which depend on whether one can assume a normal distribution and whether the number of potential sample locations is considered to be finite or infinite. Each of these methods have been or are being incorporated into VSP. A few of those that have been added to VSP will be illustrated in sections 4.2 and 4.3.

## **4.0 VSP Within Buildings Sampling Module**

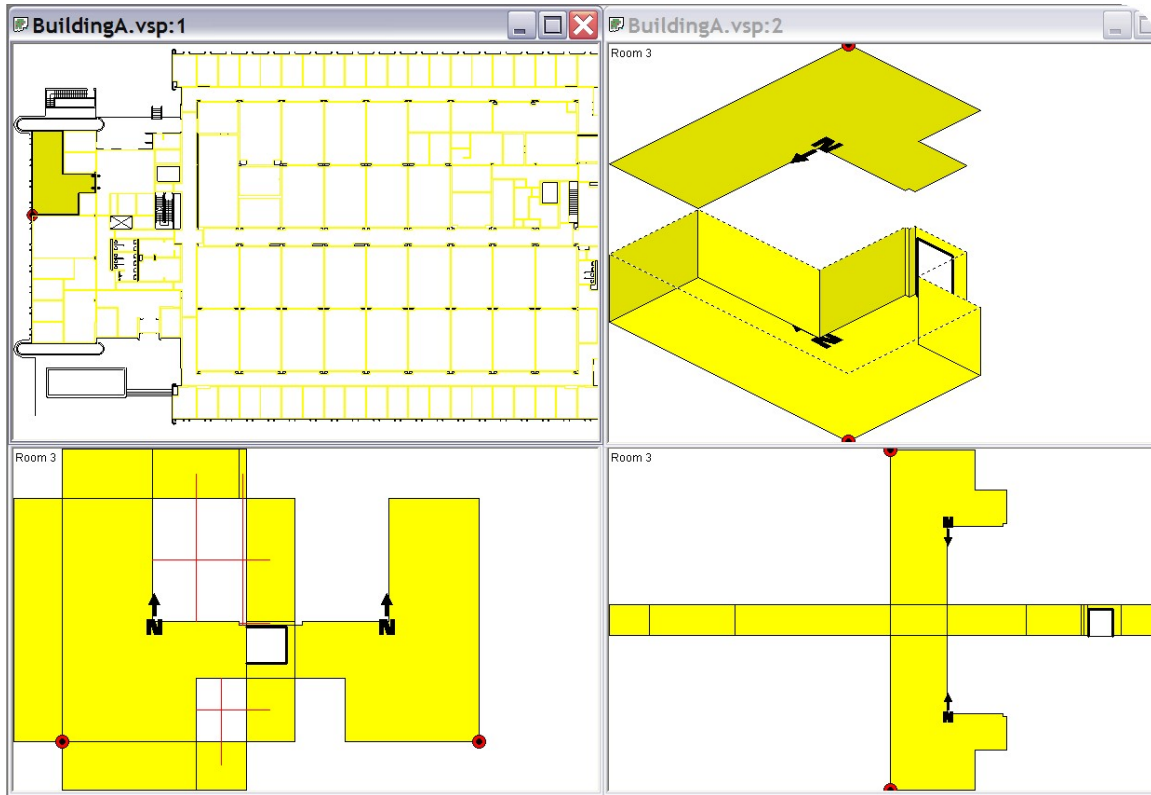
A new module has been developed within VSP to support the development of sampling designs within buildings. Additional features for visualizing and manipulating floor plans and rooms have also been added. These will be illustrated in section 4.1 and various sampling designs will be illustrated in sections 4.2 and 4.3.

### **4.1 Floor Plans, Rooms, Zones**

Using VSP tools, one can draw or import a floor plan from an existing CAD drawing. VSP accepts .dxf files or .shp files. Once a floor plan is in VSP, rooms can be defined using a room delineation tool and each room can be given an identifying label and a room height. VSP automatically calculates the room area, floor area, perimeter, and room volume. Windows and doors can be added to each room as appropriate.

Buildings and rooms can be viewed a number of different ways. Each of these views are shown in the four windows in Figure 2. The floor plan view (upper left window) is a 2-D projection viewing downward onto the floor. The perspective room view (upper right window) allows one a 3-D view of the floor, walls, and ceilings of the room highlighted in the floor plan view and permits rotations around the room viewing from different perspectives. The splayed room view (lower left window) shows the room folded out onto 2-D space. The wall-strip room view (lower right window) shows the floor and ceiling with the walls laid side by side.

User-defined zones, such as sets of rooms that are serviced by different air handling units, can also be defined and visually depicted using different colors. These visualization features greatly assist in determining and viewing the most appropriate sampling approach for various room sections, rooms, or zones within a building.



**Figure 2. Floor Plans and Room Views in VSP**

#### 4.2 Decision Rules Based on Means

VSP supports both parametric and nonparametric approaches to sampling design for sampling within buildings. Parametric approaches are appropriate if the mean or data are normally distributed. Depending on the decision rule and distributional assumption, VSP has options for simple random sampling, systematic grid sampling, collaborative sampling, or sequential sampling.

Sampling designs methodology for all three mean-based decision rules and sampling objectives outlined in Table 1 are provided in VSP. For illustrative purposes, assume one wants to develop a sampling scheme to estimate the mean contaminant concentration in 3 adjacent rooms (shown in Figure 4) and compare that mean against some regulatory threshold. Figures 3 and 4 give an example of the VSP user dialog box and the resulting output for the case when the decision rule is based on comparing a mean against a regulatory threshold. Given the DQO parameters shown in Figure 3, the required number of samples is 36. In Figure 4, the floor plan view and sample locations are shown in the upper left window, one of the rooms and sample locations are shown on the

3-D perspective room view in the window on the right, and the sample coordinates are listed in the lower left window.

**True Mean vs. Action Level**

One-Sample t-Test | Sample Placement | Costs | Data Analysis

For Help, highlight an item and press F1

Choose:

☒ True Mean  $\geq$  Action Level (Assume Site is Dirty)

☐ True Mean  $\leq$  Action Level (Assume Site is Clean)

You have chosen as a baseline to assume the site is "Dirty"

False Rejection Rate (Alpha):  %

False Acceptance Rate (Beta):  %

Lower Bound of Gray Region:  Delta = 2

Action Level:

Estimated Standard Deviation:

Minimum Number of Samples in Survey Unit: 36

☐ Use Historical

**Figure 3. VSP User Dialog Box for Comparing Mean Against Threshold**

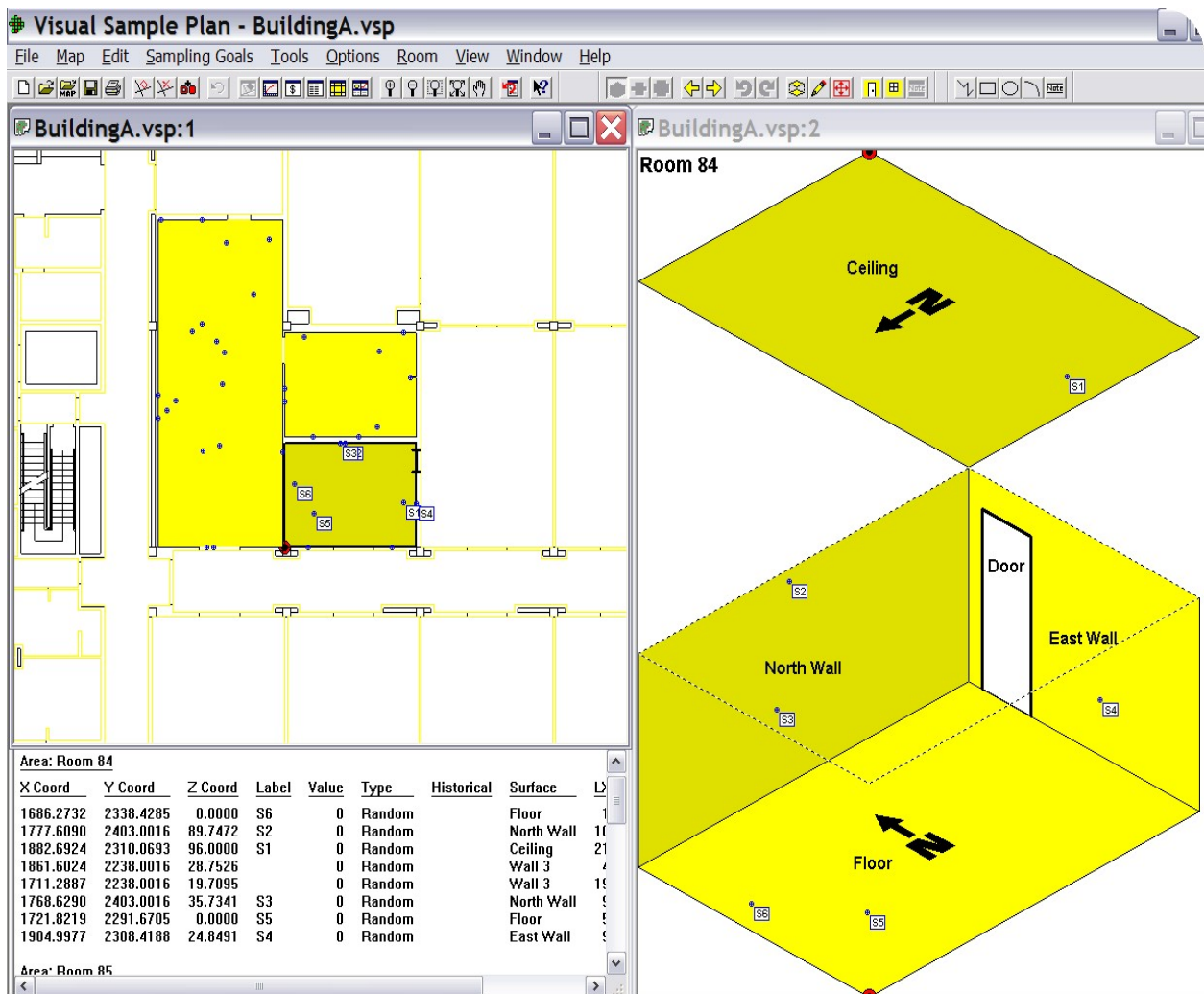
All of the usual VSP outputs including the interactive diagnostic graphics (decision performance goal diagram), sensitivity tables, online help, and automatically generated reports are produced and available for building sampling designs.

#### **4.3 Decision Rules Based on Individual Sample Results, Maximums, or Percentiles**

The four decision rules based on individual sample results, maximums, or percentiles shown in Table 1 are already available or are currently being implemented into VSP. The stakeholders should agree on which of the following statements they want to be able to make based on the sample results.

- Confident that no hotspot of a given size and shape exists.
- Confident that no more than R% of all possible measurements exceeds some limit
- Confident that none of the remaining possible measurements exceed some limit
- Confident that there are no more than X possible measurements that exceed some limit.

The sampling requirements will increase significantly as the percent or number of allowed exceedances decreases.



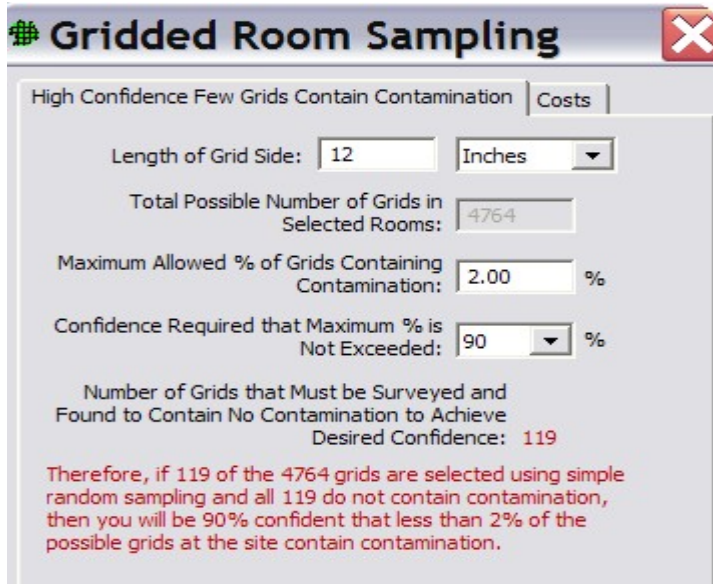
**Figure 4. VSP Output Showing Sample Locations Within Rooms and Sample Coordinates.**

The most widely recognized of these design approaches focuses on the hotspot detection decision rule. At times, one wants to ensure that if a sufficiently large contaminated area exists, enough samples are taken to have a high probability of obtaining at least one sample in that contaminated area. Therefore, if none of the samples exceeds some limit, then one can be confident that no large contaminated area exists, whereas if any individual result exceeds the limit, a hotspot may have been identified and the area is deemed contaminated. It should be noted that measurement uncertainty is assumed to be negligible.

Rather than demonstrate all design options in VSP for decisions based on individual results, maximums, or percentiles, one example will be provided herein. Suppose we are interested in demonstrating that we are 90% confident that no more than 2% of the 1x1 foot areas in three adjacent rooms are contaminated. Figures 5 and 6 show the VSP user dialog box and the selected rooms within the overall floor plan. Notice that VSP automatically calculates the number of 1x1 foot sample locations within the 3 rooms to



be 4764. VSP states that if 119 of the 4764 grids are sampled and all 119 are not contaminated, then we can be 90% confident that less than 2% of the possible 1x1 foot grids are contaminated. Figure 6 depicts the grid locations that should be sampled and the specific locations for Room 84 are shown in the splayed room view.



**Gridded Room Sampling**

High Confidence Few Grids Contain Contamination | Costs

Length of Grid Side: 12 Inches

Total Possible Number of Grids in Selected Rooms: 4764

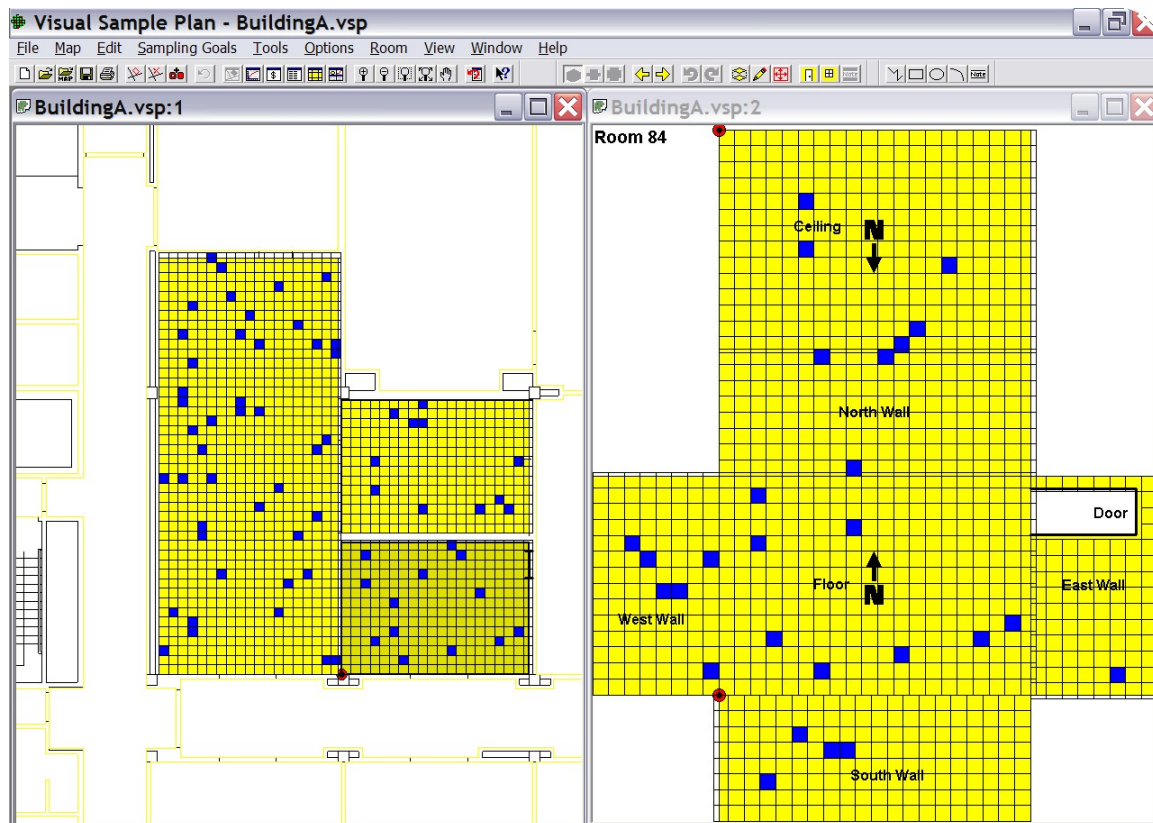
Maximum Allowed % of Grids Containing Contamination: 2.00 %

Confidence Required that Maximum % is Not Exceeded: 90 %

Number of Grids that Must be Surveyed and Found to Contain No Contamination to Achieve Desired Confidence: 119

Therefore, if 119 of the 4764 grids are selected using simple random sampling and all 119 do not contain contamination, then you will be 90% confident that less than 2% of the possible grids at the site contain contamination.

**Figure 5. VSP User Dialog Box for Percentile Decision Rule - Compliance Sampling**



**Figure 6. Gridded Room Sampling for Percentile Decision Rule.**

## 5.0 Summary

In the event of or when planning for chemical, biological, or radiological terrorist events within a building, being able to quickly develop and deploy optimal sampling strategies will be extremely important. Developing optimal within building sampling plans requires a clear determination of the decision rules and desired confidence statements. Besides selecting the decision criteria, one must determine which parameters will be estimated, what tolerance is allowed for errors in decision making, what cost constraints exist, and what practical limitations may be present. All this is facilitated by following a systematic planning process such as the DQO process.

Several statistical and judgmental sampling design methodologies have been developed and implemented within the VSP software to facilitate quick and easy development of optimal, feasible within-building sampling designs. The tools are set up to allow one to conduct tradeoff evaluations in real time to examine the effects of selected decision criteria, required confidence, and tolerance for decision errors on the number of samples required and costs. Derivation of an optimal sampling design usually is an iterative process as one works through the tradeoff evaluation process.

With the recent addition of the VSP buildings sampling module, VSP supports decisions that are based on means, individual results, percentiles, or maximums. The VSP features for setting up floor plans, rooms, and zones and visualizing in 2-D and 3-D all surfaces and sampling locations, greatly enhances one's ability to understand and communicate various sampling designs and metrics. VSP is available for free download at <http://dgo.pnl.gov/vsp>.

## 6.0 Short Bibliography

EPA. 1997. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. EPA 402-R-97-016, NUREG-1575, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 2000a. *Guidance for the Data Quality Objectives Process - QA/G-4*. EPA/600/R-96/055, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 2000b. *Guidance for Data Quality Assessment - Practical Methods for Data Analysis - QA/G-9, QA00 Update*. EPA/600/R/96/084, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 2001. *Guidance for Choosing a Sampling Design for Environmental Data Collection, QA/G-5S, Peer Review Draft*. Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

Gilbert, R. O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. John Wiley & Sons, Inc., New York.

Hassig, N. L., J. E. Wilson, R. O. Gilbert, B. A. Pulsipher. 2004. *Visual Sample Plan Version 3.0 User's Guide*. PNNL-14970. Pacific Northwest National Laboratory, Richland, Washington.

# **Incorporating Statistical Analysis for Site Assessment into a Geographic Information System**

**David K. Reichhardt, Pat Thomas**

(MSE Technology Applications; 200 Technology Way, Butte, Montana 59701, USA)

E-mail: david.reichhardt@mse-ta.com

and

**Charles S. Todd**

(Montana Tech of the University of Montana, Butte, Montana 59701, USA)

## **Abstract**

*The Fernald Closure Project (i.e., the Fernald site), a former uranium processing facility, is focusing on environmental compliance, waste management, and remediation. Since 1992, Fluor Fernald, Inc. has managed the cleanup of the 1,050-acre Fernald site for the Department of Energy and to date has remediated 672 acres of the site. A key part of this effort is certification that soils left in place following remedial efforts meet established clean-up criteria, or final remediation limits. A statistical process designed to certify that the post remediation at the Fernald site meet the final remediation limits was developed by Fluor Fernald, Inc.<sup>1</sup> The certification process is described in detail in the Fernald Environmental Management Project Sitewide Excavation Plan.*

*To streamline the statistical analysis of the soil certification data and create a system to conduct future risk analysis, MSE Technology Applications, Inc setup a geographic information system project for the Fernald site using ESRI's ArcView™ software and embedded a model of the Fernald soil certification statistical process into it. The combination of the statistical analysis and the geographic information system resulted in a useful tool, both in terms of soil certification and in evaluating future risk scenarios.*

## **Introduction**

The Fernald Closure Project (i.e., the Fernald site) (DOE, 2004) is a former uranium processing facility located 18 miles northwest of Cincinnati, Ohio. In 1989, after 37 years of operations to support the U.S. weapons program, site management shut down uranium metal production to concentrate on environmental compliance, waste management and remediation. Since 1992, Fluor Fernald, Inc., has managed the cleanup of the 1,050-acre Fernald site for the U.S. Department of Energy (DOE) and to date has remediated 672 acres of the site.

A key part of the remediation effort is the removal of contaminated soils and certification that the soils left in place meet established clean-up criteria. The clean-up criteria are known as the final remediation levels, or FRLs. A statistical process was defined<sup>1</sup> and agreed upon by the Fernald site, the U.S. Environmental Protection Agency (EPA) and the Ohio EPA to certify that the post remediation soils in a given area meet the clean-up criteria. The certification process is laid out in the Fernald Environmental Management Project (FEMP) Sitewide Excavation Plan (SEP) (DOE, 1998) and shown in Figure 1.

---

<sup>1</sup> Matthew V. Hnatov, employed by Jacobs Engineering, with input from Richard J. Abitz and Jyh-Dong Chiou, employed by Fluor Fernald Inc, developed the soil certification strategy at Fernald.

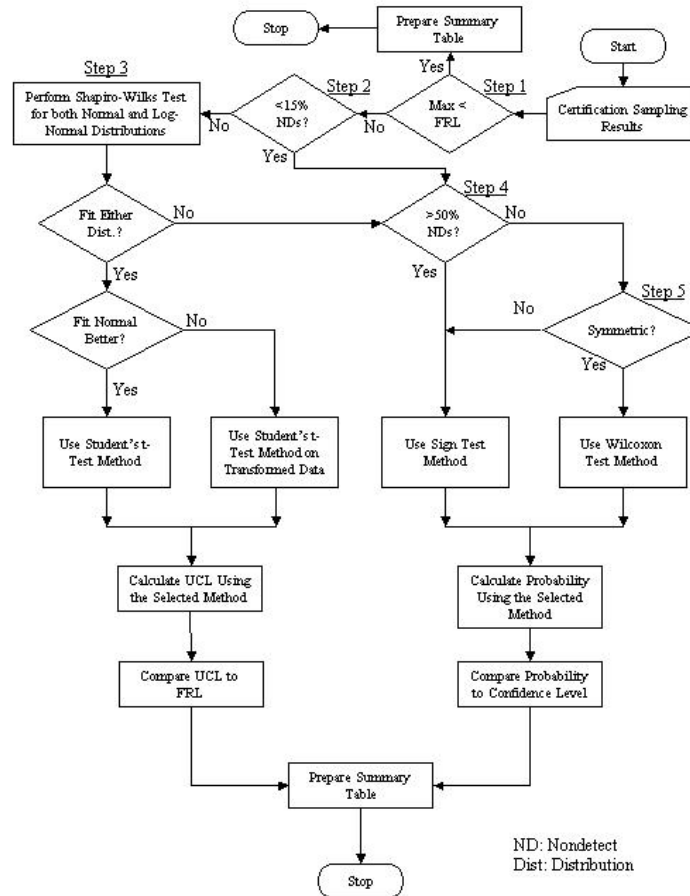


Figure 1 Fernald soil certification process flow diagram (DOE, 1998).

### Soil Remediation and Certification Process Summary

The soil remediation and certification process includes pre-excavation sampling and characterization, soil removal, post-certification soil scanning, and final soil certification sampling. Prior to the soil certification sampling, the area is divided into soil certification units. The soil certification sampling is carried out at randomly selected points within the certification units. The only criterion for the selection of the sample point location is a maximum separation distance between points determined by certification unit size and shape. For each soil certification unit, the certification sampling data are used to estimate a representative concentration for the identified constituents of concern (COC). Additionally, the probability that the concentration might exceed the FRL is also determined.

To determine if an area has met the clean-up goals, the certification data from the area must meet one of two criteria: (1) the maximum data value must be less than the FRL, or (2) the upper confidence level (UCL) of the estimated mean must be less than the FRL for normally or log-normally distributed data and the maximum value must be less than twice the FRL. If the data are not normally or log normally distributed, non-parametric statistics are used to determine the probability that the area will exceed the FRL for a given constituent. In addition to the above

criteria, the data must also pass an a posteriori test to verify that a sufficient number of values were used for the statistical tests.

This paper focuses on the incorporation of the soil certification statistical process into a geographic information system (GIS) and the application of the certification statistics through the GIS to provide a consistent and easy to use method of assessing the quality of the remediation efforts.

Because of the soil certification process design, it is applicable to data selected from multiple sampling units. Following the determination if statistical evaluation is necessary (step 1 in Figure 1, comparing the maximum data value to the FRL), the first phase of the soil certification statistics process (steps 2 and 3 in Figure 1) is the assessment of the distribution of the data. Subsequent statistics are carried out according to the nature of the distribution.

### **Incorporation of Soil Certification into GIS**

We developed a GIS project for the Fernald site using the ESRI ArcGIS™ and ArcIMS™. ArcGIS™ is a console-based system and ArcIMS™ is the web version of the software. The GIS project includes maps and aerial photographs of the site along with locations of soil certification units, certification sampling points, and certification data.

The soil certification process was incorporated into the GIS project by first preparing a standalone computer program or model of the process and then embedding this model into the GIS project. The statistics program was written using FORTRAN and compiled using Microsoft Developer Studio 97 Visual FORTRAN 5.0A. The code consists of a series of FORTRAN subroutines. Many of the subroutines are from the IMSL Statistical Library included with Microsoft Developer Studio 97. The IMSL Statistical Library is a collection of FORTRAN subroutines and functions designed and documented for use in research activities as well as by technical specialists. Additional subroutines that we used in the development of the code were taken from literature or were written based on algorithms or equations from either the FEMP SEP or referenced statistical texts. For additional details on the soil certification code see Reichhardt and Todd, 2005 and for details on the soil certification process see DOE, 1998.

### **Implementation of the GIS Soil Certification Module**

Following the development and testing of the stand-alone code for performing the soil certification statistics, we developed a module for the GIS project that would allow the user to graphically select the area for certification. The available soil certification data for the selected area is retrieved from the database and passed to the soil certification program. The program is designed to read a data table that lists the constituents, the number of samples for the particular constituent, the number of non-detect samples for each constituent data set, the FRLs for each constituent, the confidence levels for each constituent, and a list of the constituent concentrations. The results of the statistical evaluation are returned to the user and can be printed out, saved to a file, or viewed.

The following figures (Figure 2 through Figure 4) show the implementation of the soil certification module using the ArcIMS™ version.



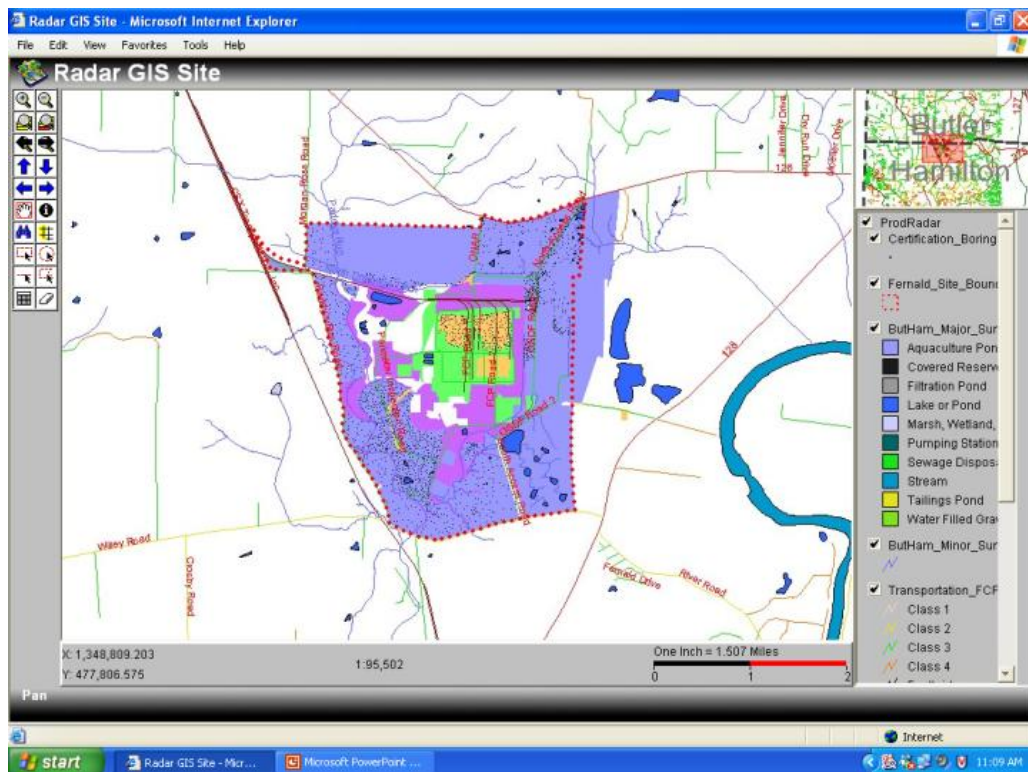


Figure 2. Screen capture of Fernald ArcIMS™ GIS project.

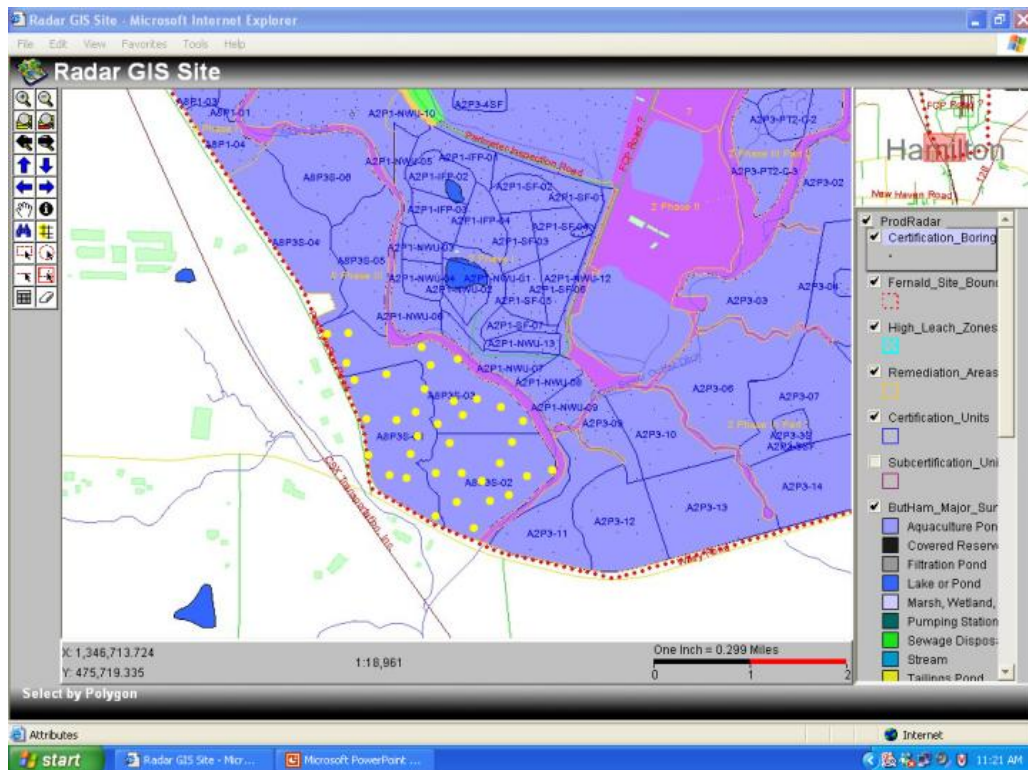


Figure 3. Screen capture of selected sample points from Fernald ArcIMS™ GIS project.

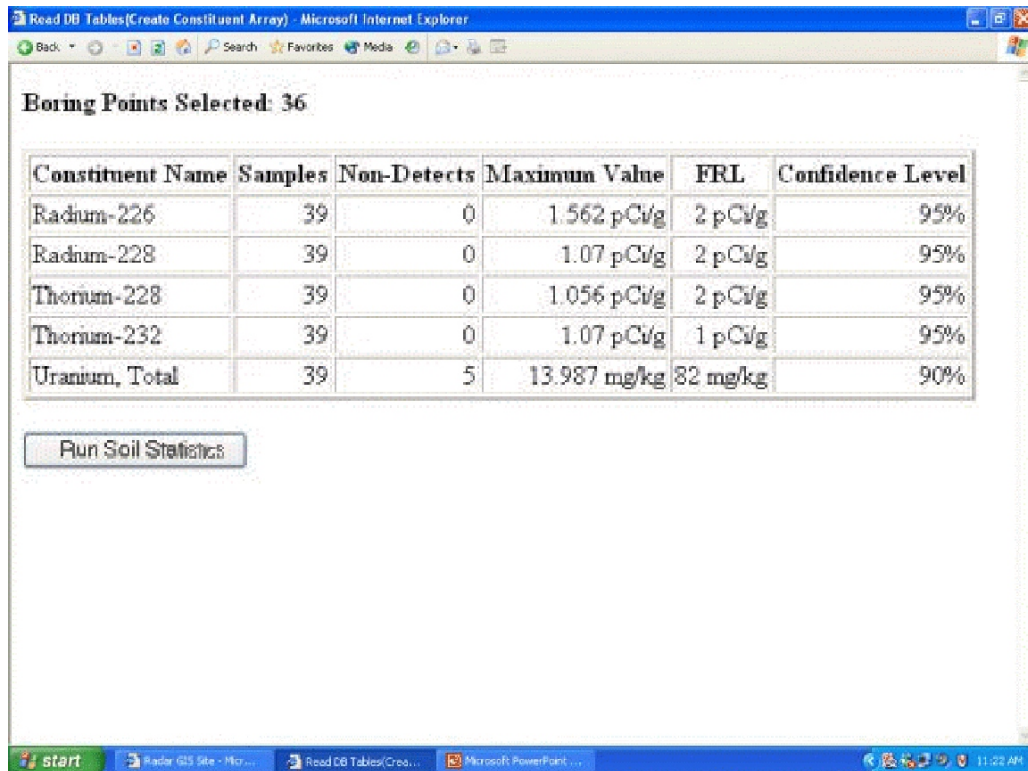


Figure 4. Soil certification data returned from Fernald ArcIMS™ GIS project.

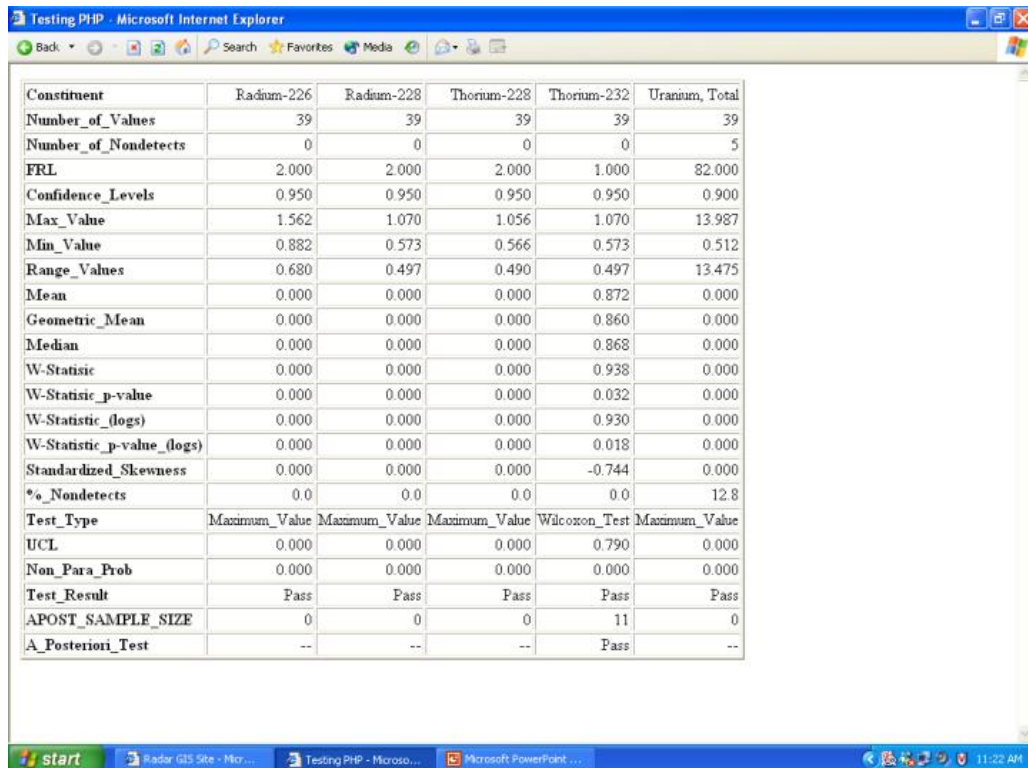


Figure 5. Results of soil certification statistics returned from ArcIMS™ Fernald project.

## **Conclusions**

The incorporation of the statistical analysis of the soil certification data into the GIS environment will allow for Fluor Fernald, Inc. project managers, regulators, and DOE stakeholders to view the soil certification results from remote locations as the data are validated and posted to the GIS project. This approach provides a consistent and statistically valid means of assessing the data. Additionally, risk analysis and “what if” scenarios may be explored using this tool. These might include assessing the clean up against more stringent FRLs, or evaluating the risk associated with different sized or shaped areas for future land use. Another important feature that may be assessed using this tool is the determination of whether or not sufficient data exist in an area for decision-making. This would apply to areas defined different from the original certification units.

## **Acknowledgements**

MSE is conducting this work through the Savannah River Operations Office at the Western Environmental Technology Office under DOE Contract Number DE-AC09-96EW96405. We have worked closely with Fluor Fernald, Inc., personnel on this effort and wish to acknowledge their help in providing information relating to the soil certification process that they developed.

## **References**

- DOE 2004. *Fernald Closure Project: About Fernald*,  
<http://www.fernald.gov/AboutFernald/aboutF.htm>, 11/17/2004.
- DOE, 1998, *Fernald Environmental Management Project (FEMP) Sitewide Excavation Plan (SEP)*, 2500-WP-0028, Revision 0. Fernald Area Office, Ohio.
- Reichhardt, D.K., C. S. Todd, 2005, *A soil certification statistics process implemented using a geographic information system*. Proceedings 2005 Spring Simulation Conference.



## **The Office of Pesticide Programs' Pesticide Data Program Environmental Indicator Project**

Philip Villanueva, Office of Pesticide Programs  
US EPA

The EPA's Office of Pesticide Programs' (OPP) mission is to protect human health and the environment from unreasonable adverse effects from the use of pesticides. In support of that mission, the Office may at times take mitigatory actions such as cancellation, reduction in use rates, and/or extension of pre-harvest intervals as a part of its risk management process. Given the increasing interest in quantitative measures of the effects of environmental programs on public health and the environment on the part of OMB, the Administrator, and Congress, it is important for OPP to be able to measure the effects of such regulatory actions. As stated in the Administrator's State of the Environment Report:

*Environmental indicators are measures that track environmental conditions over time... [And] ... focus on outcomes—actual environmental results, such as cleaner air and water or improved human health or ecosystem condition—rather than on administrative actions, such as the number of permits issued. At one time, administrative measures of performance were considered sufficient indicators of progress. While administrative measures track what actions have been taken, they don't tell us whether those actions actually improved the environment or human health. Understanding the effectiveness of environmental programs, and measuring actual progress, requires indicators of health and environmental conditions.*

OPP has available a variety of data sources which can be used to develop environmental indicators. One such source is the USDA's Pesticide Data Program (PDP), which can be used to track trends in pesticide concentrations in a variety of produce and other food commodities. PDP provides a reliable source of pesticide monitoring data on a broad cross-section of commodities selected to represent the food supply available in the US. Since approximately 1992, an extensive, statistically-based national dataset on pesticide residues has been collected by PDP on a regular, continuing basis. This data is collected in a manner designed to simulate as closely as possible "dinner-plate" exposures to consumers and is routinely used in support of OPP's registration, re-registration, and risk management decisions. Over the years, hundreds of thousands of analyses have been performed on tens of thousands of samples of fruits, vegetables, processed foods, grains, meats, and dairy products for over a hundred pesticides and their metabolites. Given the extensive data available for any individual crop-pesticide combination, the USDA's PDP program can provide a valuable source of data for an environmental indicator and, appropriately analyzed, an excellent quantitative metric for pesticide residues in produce and other food commodities. OPP is continuing to develop this data source as an indicator metric and this paper provides a brief introduction to the work being done and summarizes some of the lessons learned.

## **The USDA Pesticide Data Program**

The USDA Pesticide Data Program (PDP) is designed to provide on a continuing basis extensive, statistically-representative data on pesticide residues in foods. The statistical design of the sampling program permits scientific statements to be made about the distribution of certain pesticide residues in selected commodities. The program was initiated, in part, as a result of the USDA's charge to collect data on pesticide residues in food. The 1993 release of the NAS report "Pesticides in the Diets of Infants and Children" and the 1996 Congressional passage of the Food Quality Protection Act and its requirement that USDA provide data to support US EPA implementing this act re-enforced the need for and importance of a continuing, statistically-based residue data collection program of national scope. PDP generates data that are statistically sound, publicly available, extensively quality-assured, and widely disseminated. The PDP data provided by USDA has developed into a critically important component of virtually all of OPP's risk assessments is a critical component of sound, scientifically-based risk-management decisions regarding pesticide registrations and mitigation actions.

Importantly for an indicator metric, the PDP sample collection procedures are specifically designed to produce dietary exposure estimates that closely approximate real world exposures. Samples are collected by USDA at terminal markets and warehouses immediately before these commodities are shipped to supermarkets and other retail establishments. Sampling is conducted in accordance with a statistically-based plan designed and put in place by USDA's National Agricultural Statistics Service (NASS) to be representative of the US food supply. Samples are prepared by the analytical laboratory as if for consumption (i.e., they are washed, peeled, and/or cored, as appropriate) and thus are more reflective of actual exposures than data typically available from field trials or FDA monitoring programs. Thus measurements simulate as closely as possible, dinner-plate exposures to consumers. The PDP data is available in downloadable electronic format from the USDA PDP website (<http://www.ams.usda.gov/science/pdp/>) and can be easily transferred, manipulated, imported, analyzed, and summarized. The PDP data is also available in the form of a CD-ROM with a search utility included.

## **Statistical Questions and Issues**

OPP formulated three initial environmental indicator questions, anticipating that these questions could be answered by appropriately analyzing the PDP data. To ensure such questions *could* be addressed, a particular crop-pesticide combination would have to meet certain criteria. One of these criteria was that a minimum percent of the analytical analyses performed by the PDP laboratories for a particular crop-pesticide combination resulted in detectable residues. Another obvious criterion was that multiple years of PDP data be available for the crop-pesticide combination of interest. Once a crop-pesticide combination had been identified as a useful dataset, the following environmental indicator-related, statistical questions were investigated:

- 1. Are the mean concentrations of pesticides on domestically-produced treated commodities changing over time?*
- 2. Are the distributions of pesticide concentrations on domestically-produced treated commodities as a whole shifting over time?*

*3. Are “percent detects” of pesticides on commodities changing over time ? Do factors such as origin, season, and region affect (or account for) these temporal changes or trends?*

During our initial investigation of these issues, we found that a number of technical issues arose and standard procedures and statistical protocols needed to be developed. These are more fully discussed below as they relate to each of the questions of interest.

*Are the mean concentrations of pesticides on domestically-produced treated commodities changing over time?*

Perhaps the most direct and obvious technique for investigating temporal trends in residue concentrations is to estimate average (arithmetic mean) concentrations on a year-by-year basis. Estimating mean concentrations on a year-by-year basis in a manner that is meaningfully representative of trends can, however, be problematic. For example:

- Arithmetic means can be profoundly affected by a few outlying values. Thus, the arithmetic mean is considered an unstable value that is often driven by characteristics of the tails of the distribution.
- In many cases, a sizable fraction of PDP residues are non-detectable. While many of these represent “true zeroes” in the sense that the specific commodity item being analyzed was not treated, at least some of these “ND” reports represent values that *are* present but at concentrations lower than the concentrations that are reliably detectable by PDP’s analytical methods. In addition, for those samples which *do* contain detectable residues, but at concentrations less than the PDP’s limit of quantitation (LOQ), PDP reports these concentrations, by default, as  $\frac{1}{2}$  the LOQ. Thus, an ND can represent a true zero level (if the commodity was not treated) or a concentration between 0 and the LOD. In addition, a concentration present between the LOD and the LOQ will be represented by a default value of  $\frac{1}{2}$  the LOQ.<sup>1</sup> Thus, a sizable percentage of the reported pesticide residue concentrations in the PDP database are default values and are represented by zeroes (for values <LOD) or  $\frac{1}{2}$  LOQ (for values between the LOD and LOQ).
- PDP analytical methods are improving over time, and limits of detection and limits of quantitation in later years are often lower -- sometimes significantly so -- than those in earlier years. This may create trends which are artifacts, due more to improved analytical techniques than to any real changes in residue concentrations. This is also a case where

---

<sup>1</sup> Substitution methods (e.g., replacing ND values with  $\frac{1}{2}$  LOD or  $1/\sqrt{2}$  LOD values) are often used in this situation but can dramatically, and inappropriately, affect calculated mean concentrations. Alternatively, one can estimate a potential range of mean concentrations for each pesticide-crop-year combination by alternately making worst case and best case assumptions. More specifically, one could estimate the lower bound on an “average” concentration by assuming that all < LOD concentrations are zero and that all  $\frac{1}{2}$  LOQ’s are at the LOD; an upper bound on this average concentration could be estimated by then subsequently assuming that all <LOD are present at the (full) LOD and all  $\frac{1}{2}$  LOQ default values are present at the (full) LOQ. However, with a large percentage of the dataset consisting of <LOQ and <LOD values, the bounds can be very wide and would not be informative with respect to trends in average concentrations.

simple substitution methods (see footnote) which are often typically used can lead to gross errors in assessing and evaluating trends.

In order to avoid or minimize these problems, OPP designed its residue estimation methods as follows:

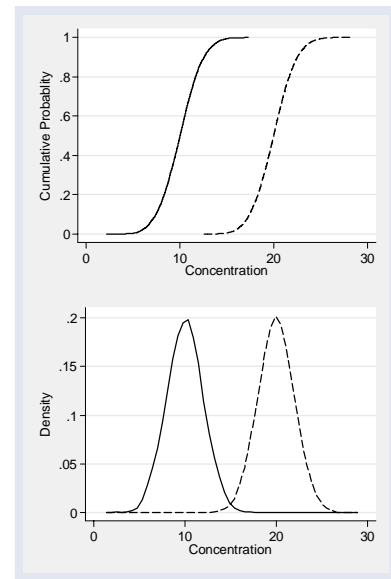
- For a skewed or lognormal distribution, the geometric mean (as opposed to the arithmetic mean) is generally considered to be more representative of the “middle” or bulk of the distribution and thus might be more reflective of a “typical” value. It is also less subject to influence by outliers. In order to better reflect this typical concentration, trends were evaluated using geometric mean concentrations. In addition, geometric means have better statistical properties than arithmetic
- OPP used (domestic) percent crop treated data available from USDA’s National Agricultural Statistics Service (NASS) and other databases to remove those “ND” concentration values which represent crops that were not treated. As a specific example, if 30% of the domestic crop in any given year was treated (and 70% was not treated) and 80% of the residue values were ND, 70% of the domestic residue values were removed from the data set since they represent untreated non-detects and 10% represent treated commodities (with non-zero residues) that are <LOD. Removing these “true zero” data from the dataset provides for a better representation of residue trends in treated crops that appropriately discounts the artificial year-to-year variations in percent crop treated. Removing these “true zero” data points permits a more robust analysis of the remaining (treated commodity) information that better -- and more directly-- reflects application practices rather than a combination of percent crop treated, market forces, pest pressures, etc. that are not controlled or influenced by EPA mitigation actions.
- For the remaining residue values (all of which were assumed to contain some residue concentration greater than zero since all were treated with pesticide), maximum likelihood estimation (MLE) methods were used to assign specific concentration values to those concentrations that were reported by PDP under their sample reporting protocol as <LOD and ½ LOQ. When the functional form of a probability distribution is known (e.g, normal, lognormal, etc.), measurements associated with <LOD values can be more appropriately and accurately assigned using MLE methodology. MLE methods select a probability density function that maximizes the likelihood of observing the collected data. The assigned values obtained through MLE methods will be the most consistent with the observed sample data (both actual measurements and <LOD data). Additional procedures can then be used to assign values to the censored data based on the “most likely” set of parameters given the data actually observed.

Using these procedures, OPP believes that it has generated a reliable metric to evaluate and judge temporal trends of residue concentrations.

Are the distributions of pesticide concentrations on domestically-produced treated commodities as a whole shifting over time?

While the analysis of temporal trends in geometric mean concentrations (described above) is valuable and appropriately reflects “typical” or “representative” concentrations in treated domestic produce, it is also useful to evaluate changes or shifts in the shape or location of the distribution as a whole. If, over time, the distribution shifts toward lower concentrations, this would indicate that residue concentrations are becoming smaller overall. Graphical plots can be particularly effective in describing these locational shifts and in conveying important aspects of any conclusions.

An (artificial) example of such a shift is shown in the figure to the right. As can be seen, the distribution represented by the solid line is universally lower than that represented by the dashed line. If the solid line represents residue concentrations during a later period and the dashed line represents concentrations during an earlier period, then this suggests an overall “shift” in the distribution over time that can be readily and easily evaluated using statistical measures. Analyses can be conducted using parametric and/or non-parametric tests, as appropriate, to permit statistically-supported statements regarding differences or shifts in pesticide residue concentrations over time to be made.



Are “percent detects” of pesticides on commodities changing over time? Do factors such as origin, season, and region affect (or account for) these temporal changes or trends?

While the previous two questions consider the changes in residue *concentrations* in treated commodities of domestic origin over time, this question deals with the changes in the *proportions* of any given commodity that have “detectable” residues over time. Specifically, this issue is concerned with whether the proportions of a given commodity with residues above some detectable level are increasing, decreasing, or remaining constant over time. This issue can be readily addressed through tabular summaries, bar graphs, and trend tests (e.g., Cochran-Armitage trend test).

In determining whether a trend in the percentage of a given commodity containing detectable residues exists and in assigning, evaluating, or ascribing potential reasons for these changes, it is critical to consider whether these putative differences exist as a result of confounding factors.<sup>2</sup>

---

<sup>2</sup> Frequently, this document will refer to “percent detects”. Since the detection limit can change (generally decreases) with time and this can lead to an artificial trend, this would be more precisely stated as the percentage of samples which contain residues that exceed some operationally-defined threshold, or Percent above Maximum Limit of Detection (Percent above MaxLOD)

More specifically (and as an example), it may be that overall percent detects are increasing with time despite the fact that mitigation actions have been taken. Closer examination of the data might reveal, however, that these increases can be attributed to imported produce – with domestic produce actually showing declines in the percentage of samples with detectable residues.<sup>3</sup> If *confounding* effects or *effect modification* (aka “interaction”) are not considered, incorrect inferences can be drawn with respect to trends and their sources.

OPP chose to use logistic regression and accompanying odds ratios to evaluate the trends in “percent detects.” If the response is dichotomous, logistic regression is a modeling strategy that relates the logit of a binary event to a variety of potentially explanatory variables. An important benefit of logistic regression as a modeling strategy is that odds ratios can be derived from the logistic regression coefficients. These odds ratios appropriately consider, adjust, or otherwise account for confounding and effect modification (if present). Judgments regarding the influence (or lack of influence) of specific variables can accordingly be made, *in a manner which implicitly holds constant all other variables*. The specific variables considered for potential introduction and inclusion in the base model are

- Origin (domestic vs. import)
- Year (as a continuous variable)
- Geographic region of US (Northeast, South, Midwest, or West)
- Season (spring, summer, fall, or winter)

Due to differences in LOD levels among and within labs and over time, “presence” and “absence” of detectable levels were dichotomized into “presence above some maximal LOD” or “absence above some maximal LOD,” where the maximal LOD will generally represent a high-end LOD for the commodity-pesticide pair of interest.

## Lessons Learned

- Estimating mean or geometric mean concentrations for datasets with a large number of non-detects (i.e. censored values) can be difficult. MLE-base methods (such as the “robust method” described by Helsel<sup>4</sup>) are frequently the most appropriate approach and are far better than standard LOD/2 imputation techniques.
- Indicators associated with concentration measured over time can be problematic in that limits of detection frequently improve. In such cases, binary classification (e.g., detected, not detected) can be misleading unless improvements in analytical techniques are considered. For such

---

<sup>3</sup> In fact, as an extreme case it is possible to find that percent detects is increasing over time, but – looked at separately – percent detects are decreasing in both imported and domestic produce. This condition, known in statistical theory as “Simpson’s Paradox”, can arise if (for example) there are a greater percentage of detects in foreign fresh produce than in domestic fresh produce BUT foreign fresh produce represents an increasing percentage share of the U.S. market over time.

<sup>4</sup> Helsel, Dennis R. (1990). “Less than obvious: Statistical treatment of data below the detection limit,” *Environmental Science & Technology*, 24, No. 12, 1766-1774.

dichotomous data, logistic regression is an approach which allows the statistician to appropriately consider confounding effects.

- Developing indicators is a delicate balancing act between the desire on the part of program management to develop simple, easy-to-understand metrics and the desire on the part of the statistician to avoid misleading the audience. Graphics and footnoted summary tables can be particularly effective and satisfy both management and technical needs.